BMJ Open

Accuracy of Mortality Statistics in Palestine: A retrospective cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-026640
Article Type:	Research
Date Submitted by the Author:	12-Sep-2018
Complete List of Authors:	Massad, Salwa; Palestinian National Institute of Public Health, Johansson, Lars; Uppsala University, Department of Public Health and Caring Sciences Ramlawi, Asad; Ministry of Health Dalloul, Hadil; Palestinian National Institute of public Health Rayyan, Izzat; Palestinian National Institute of Public Health Salman, Rand; Palestinian National Institute of Public Health
Keywords:	EPIDEMIOLOGY, Protocols & guidelines < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, PUBLIC HEALTH

Title Page

Accuracy of Mortality Statistics in Palestine: A retrospective cohort study

Salwa G Massad1, Lars Age Johansson2, Asad Ramlawi 3, Hadil Dalloul1, Izzat Rayyan1, Rand Salman 1*

[1]Palestinian National Institute of Public Health, Ramallah, Palestine, [2]Department of Public Health and Caring Sciences, Uppsala University, Uppsala, Sweden, [3] Palestinian Ministry of Health, Ramallah, Palestine.

*Correspondence to: Salwa G Massad, PhD. Research manager, Palestinian National Institute of Public Health, Hay El Razy Ramallah, West Bank/Palestine Tel: 00970598161071, Email: salwamassad@gmail.com

Email addresses:

Salwa G Massad (salwamassad@gmail.com)

Lars Age Johansson (lars.age@bredband.net)

Asad Ramlawi (ramlawi_asad@hotmail.com)

Hadil Dalloul (ramlawi_asad@hotmail.com)

Izzat Rayyan (rayyan67@yahoo.com)

Rand Salman (salmanr@who.int)

Key words: cohort study, hospital based study, health system evaluation

Word count 4795

Abbreviations:

CoDR: Cause of Death Registry

DNF: Death notification form

EDC: Medical Extraction Forms

ICD: International Classification of Diseases

MoH: Ministry of Health

MoI: Ministry of Interior

PHC: Primary Health Care

PHIC: Palestinian Health Information Centre

Accuracy of Mortality Statistics in Palestine: A retrospective cohort study

Abstract

Objective To examine the accuracy of mortality statistics in Palestine, to identify gaps, and to monitor and evaluate planned interventions aimed at improving the Statistics.

Study Design and Setting We retrieved the medical records for a random sample of hospital deaths reported in 2012: 371 deaths in the West Bank and 199 deaths in the Gaza Strip.

Results Data in the Palestinian Health Information Center Registry had a low degree of accuracy: less than half of the underlying causes stated the correct cause of death. The two major reasons are: 1) Discrepancies between the coding and classification procedures set out in the International Classification of Diseases (ICD) and the procedures applied by the Palestinian Health Information Center (PHIC); and 2) Incomplete or otherwise unsatisfactory reporting on a considerable number of death notification forms (DNFs).

Conclusion Procedures for coding and classification at the PHIC deviate considerably from the international norms defined in the ICD and account to a considerable extent for the discrepancies between the cause of death determined on the medical data on the death extracted from the deceased patient's hospital records and the cause of death coded by the PHIC. We recommend the introduction of international coding software for coding and classification, and a review to improve data handling in hospitals, especially those with electronic patient records.

Summary box

Strengths and limitations of this study

- The present study is the first assessment study of accuracy of mortality statistics in Palestine and to our knowledge, it is the first assessment study in the Region.
- The present study is the first one to examine the completeness and accuracy of death notification forms.
- The study was limited to hospital deaths.
- The presence of deaths that occurred several decades ago in the sample indicates that reporting is incomplete. It is important to estimate the number of non-reported deaths and determine if the causes of death in such cases differ significantly from those actually reported to the authorities.

INTRODUCTION

Statistics on causes of death are essential indicators of the overall health and quality of life of a population. They can be used to monitor the status of health, estimate the burden of major diseases over time and across geographical regions, and to assess the impact of health interventions. However, this requires accurate statements on the cause of death and accurate statistics collected in accordance with international standards (1). We carried out this study to examine the accuracy of mortality statistics in Palestine. It will serve to identify information gaps, and to monitor and evaluate planned interventions for improving mortality data in Palestine.

In Palestine, the Cause of Death Registry (CoDR) held by the Palestinian Health Information Centre (PHIC) at the Ministry of Health (MoH) is part of the national vital statistics database and holds information on deaths and the causes of deaths since 1994. Prior to that date, causes of death were registered with the Israeli Civil Administration.

The main source of data for the CoDR is the Death Notification Form (DNF), issued by the attending physician. The DNF includes information about the main and underlying causes of death. The DNF is given to the family of the deceased, who are the ones responsible for notifying the PHC of the death. In the Gaza Strip, some hospitals keep a copy of the DNF but others do not. The PHC or the family of the deceased sends the DNF to the Ministry of Interior (MoI). At the end of each month the PHC Directorate compiles all DNFs and mails them to the PHIC for coding and registration. The PHIC also receives some DNFs from governmental hospitals. There are no DNFs for neonatal deaths; instead, a monthly report of neonatal deaths is

sent by each hospital to the PHIC. The PHIC codes the underlying causes of death and enters the death into the database.

METHODOLOGY

Study design

This retrospective death registry-based study examined a stratified random sample of DNFs of patients who died in hospitals in the West Bank and Gaza Strip in 2012. We randomly selected 600 deceased from the Cause of Death Registry: 400 from the West Bank and 200 from the Gaza Strip. The sample size was estimated using the SampleXS program (2) and based on the following assumptions: the number of registered deaths in 2012 was 12,000; and the completeness of death certificates was 60% based on the recent PCBS study of the completeness of the Cause of Death Registry (3); the margin of error 5%; and the design effect 1.5.

About 70% of deaths in the West Bank occurred in government hospitals, while in the Gaza Strip all deaths were in government hospitals. We took a random sample of hospital deaths from the following four age groups: less than 1 year, 1-17 years, 18-64 years and 65+ years. The study was carried out in 2014 in close collaboration with the Ministries of Health (MoH) in the West Bank and Gaza Strip. The response rate was 100%: all selected hospitals in the West Bank and Gaza Strip agreed to participate in the study.

Material

We used several sources of data to assess the accuracy and completeness of the mortality data. For each death in the sample we collected the following data:

(i) The Death Notification Form (DNF)

- Scanned image of the DNF
- Formal assessment of the DNF, such as legibility of the handwriting and if the administrative data had been filled out correctly with the certifier's signature.
- The causes of death in free text and administrative information were extracted from the DNF.

(ii) Data from PHIC:

- Date on which the death was registered and the ICD codes for the underlying causes of death as recorded by the PHIC.
- ICD codes for underlying causes of death as recorded by the PHIC in the Death Registry.

(iii) Medical Extraction Forms (EDC):

Medical data on the death extracted from the deceased patient's hospital records. A team of specially trained physicians from each hospital in the study extracted data from the hospital record on the train of events leading to the death. The extraction file contained data on the reason for hospitalization, previous medical history and the course of events during hospitalization.

(iv) Hospital case summaries:

- Scanned image of the original hospital case summary.
- Case summary based on hospital records, completed by specially trained physicians in each hospital.

Figures 1 and 2 summarize data availability and attrition in the West Bank and Gaza Strip. In the West Bank, hospital records could be found for 371 cases of the 400 randomly selected deaths. In the Gaza Strip, we found hospital records for 199 of the 200 randomly

selected deaths. The PHIC Registry gave an ICD code for the underlying cause of death in 358 of the 371 cases in the West Bank sample for which hospital records were available. In the Gaza Strip, 189 of the 199 cases in the sample had an ICD code for the underlying cause of death and 70 cases also included multiple causes of death registered by the PHIC. No multiple causes were present in the West Bank sample. Scanned images of the DNFs were available for 320 deaths in the West Bank sample. EDC data on file were available for 365 West Bank deaths and for all Gaza Strip deaths. Table 1 summarizes the characteristics of the data.

Table 1. Data characteristics of the records available for the study

	N (%)
West Bank	**371**
- Governmental hospitals	265 (71%)
- Hospitals using electronic patient records	85 (23%)
- Age	
< 1 year	75 (20%)
1-17 years	45 (12%)
18-64 years	149 (40%)
65+	102 (27%)
- Gender	
Male	200 (54%)
Female	171 (46%)
Gaza	**199**
- Governmental hospitals	199(100%)
- Hospitals using electronic patient records	0 (0%)
- Age	
< 1 year	54 (27%)
1-17 years	19 (10%)
18-64 years	72 (36%)
65+	54 (27%)
- Gender	
Male	99 (50%)
Female	100 (50%)

Figure 1: Data availability and attrition in West Bank

Figure 2: Data availability and attrition in Gaza Strip

STATISTICAL ANALYSIS

Completeness and timeliness

We examined the proportion of complete DNFs in terms of personal data, immediate and underlying causes of death, and the attending physician and notifying person. We also examined the timeliness of registration, measured from the time of death to the time the deceased was registered in the Cause of Death Registry.

Underlying cause of death based on the patient's hospital file and on the DNF

Using the EDC, we determined the underlying cause of death and coded it using ICD-10. All coding, both of the original DNFs and of the EDC data, was done using Iris coding software (4). The version used (V4) corresponds to the 2013 edition of ICD-10. Iris fetches an ICD code s for each diagnostic expression from a dictionary. Next, Iris prepares an input string of ICD codes for the ACME module (developed by the US National Center for Health Statistics). Finally, ACME selects an underlying cause according to the ICD rules and guidelines. After review and coding, we compared the three underlying causes – the one coded by PHIC; the one stated on the DNF form, but coded and classified by Iris; and the one according to the EDC – pair by pair:

- *Underlying cause on the DNF, coded by Iris vs Underlying cause in the PHIC Registry*

This comparison measures PHIC compliance with international rules for selection of the underlying cause of death.

- *Underlying cause on the DNF, coded by Iris vs Underlying cause according to the EDC*

This measures the accuracy of the original DNF.

- *Underlying cause according to the EDC vs Underlying cause in the PHIC Registry*

This comparison measures the overall accuracy of the underlying cause registered in the PHIC Registry.

We compared the underlying causes at two levels of detail: ICD detailed level (henceforth referred to as ICD 4-character level) and ICD block level. The principal use of measurements at detailed level data is to estimate the precision of the coding and classification, while measurements at ICD block level provide a broad assessment of the general trustworthiness for public health purposes.

For cross tabulation, we examined the accuracy of mortality statistics by gender and age group of the deceased patient, hospital affiliation (governmental and non-governmental), type of hospital records and underlying cause of death using chi square test. Tests of significance were two-sided with p- value ≤0.05. We divided the underlying causes according to the EDC into eight main diagnostic groups based on the frequency of the most common underlying causes according to the EDCs.

- Neoplasms (ICD-10 Chapter II)
- Metabolic diseases (ICD-10 Chapter IV)
- Cardiac diseases (ICD-10 I00-I51)
- Cerebrovascular diseases (ICD-10 I60-I69)
- Perinatal conditions (ICD-10 Chapter XVI)
- Congenital anomalies (ICD-10 Chapter XVII)
- External causes (ICD-10 Chapter XX)
- Other causes of death (ICD-10 chapters and codes not included elsewhere). ("Other" includes infectious diseases, diseases of the blood, neurological diseases, diseases affecting vision and hearing, respiratory diseases, gastrointestinal diseases, skin

diseases, musculoskeletal diseases, urogenital diseases and symptoms with no clear connection to a single underlying disease.)

Selecting the underlying cause of death

The classification expert (LAJ) checked the diagnoses registered in the DNF files against the scanned images of the DNFs and corrected the registered text if it differed from the scanned image. Similarly, LAJ checked the causes of death reported on the EDCs against the case summaries and against the scanned images of the hospital summaries. If an EDC was inconsistent with the case summary or the hospital summary, the EDC was corrected according to the description in the summaries.

Also, for each DNF, the following markers were used to record non-compliance with WHO instructions for completion of the death certificate: abbreviations used in the medical section, illegible writing in the medical section, sequencing errors in Part 1, symptomatic or secondary condition reported as underlying cause of death, cause of death insufficiently specified, and other reporting errors (such as reporting several causes on the lowest completed line in Part 1, placing the underlying cause in Part 2 of the certificate, or using the wrong field on the DNF to report causes of death).

Patient and Public Involvement

Not applicable for this study as we used mortality statistics and data was de-identified prior data analysis.

RESULTS

Completeness of DNF forms and other characteristics

Table 2 gives a summary of other characteristics of DNFs unrelated to the accuracy of the underlying cause of death. Although most deaths had been reported within a year (median of 89 days; range excluding extreme values 0-365 days), occasionally deaths from the 1980s and 1990s were only reported to the authorities in 2012. A fairly high number of DNFs (23% West Bank, 26% Gaza) reported an underlying cause of death that presumably developed as a complication of some other condition, yet the underlying condition is not recorded. For example, in several DNFs, kidney failure was reported as the underlying cause of death, but there was no mention of the reason why kidney failure developed, such as diabetes, glomerulonephritis or urinary obstruction. Administrative data (Part 1 of the DNF) were complete in only 7% and 2% of DNFs from hospitals with electronic patient records and paper patient records respectively. Other problems with electronic records included several deaths that had been registered twice or more, and some of the deceased recorded in the hospital files were still alive.

Table 2. Completeness and other DNF characteristics

	West Bank (N=320)		Gaza Strip (N=199)	
	cases	%	cases	%
No medical data on DNF	23	7.2	1	0.5
- electronic records (71)	15	21.1	-	-
- paper records (249)	8	3.2	1	0.5
Administrative data complete (Part I)	9	2.8	108	54.3
- electronic records (71)	5	7.0	-	-
- paper records (249)	4	1.6	-	-
Abbreviations used	132	41.2	76	38.2
Illegible writing	96	30.0	11	5.5
Sequence errors	64	20.0	38	19.1
Non-informative UC	74	23.1	51	25.6
UC lacking specificity	36	11.2	12	6.0
Irrelevant information	105	32.8	22	11.0
Other certification errors	92	28.8	40	20.1

Competing causes of death	36	11.2	23	11.6
Certifier's signature complete	159	59.7	50	25.1
Registry delay (days; median)	89		57	

Agreement between the three underlying causes (Table 3)

DNF- EDC agreement

Agreement on the underlying cause was highest between the underlying cause derived from the original DNF and the medical extraction forms (DNF-EDC). At the most detailed level, the agreement was 56% for the West Bank sample and 52% for the Gaza Strip sample. The difference between the two sample groups was not statistically significant.

DNF-PHIC agreement

DNF-PHIC agreement between the DNF coded in line with international standards and the underlying cause as coded by the PHIC was 23% in the West Bank sample at ICD 4-character level and 39% in the Gaza sample. The difference between the samples is statistically significant. At ICD block level, the West Bank sample had a DNF-PHIC agreement of 46% and the Gaza sample 53%, a difference that is not statistically significant.

EDC-PHIC agreement

Agreement was lowest between the EDC and the underlying cause recorded in the PHIC Registry. At ICD 4-character level, EDC-PHIC agreement was 19% in the West Bank sample and 31% in the Gaza sample.

Table 3. Agreement on underlying cause by ICD level: point estimates, and for ICD-4 character level and ICD block level and 95% confidence intervals

		ICD 4-character		ICD block level	
		PE (%)	95% CI	PE (%)	95% CI
DNF-EDC	West Bank	56	51-62	68	63-73
	Gaza Strip	52	44-59	62	55-68
DNF-PHIC	West Bank	23	19-28	46	41-51
	Gaza Strip	39	31-46	53	46-60
EDC-PHIC	West Bank	19	15-23	44	38-49
	Gaza Strip	31	24-38	44	37-51

Agreement by sex, age, type of hospital and type of deceased hospital records

In both samples, there was no noticeable relationship between the sex and age of the decedent and the accuracy of the cause of death. There were few statistically significant differences between males and females in underlying cause agreement (21% vs.16%, respectively). Also, there were few significant differences between the four age groups, either at ICD 4-character or ICD block level. At a detailed level there was, however, significantly lower agreement for the youngest age groups in the two comparisons involving the codes in the PHIC Registry (DNF-PHIC and EDC-PHIC) (7% and 8%) compared to 27% and 24% at age group 18-64years.. At block level, only the DNF-PHIC comparison showed a significant difference between children <1 year and older people >64 (43% vs 60% respectively), probably reflecting difficulties in the coding of congenital anomalies and perinatal causes.

Also, the West Bank sample showed only minor differences between governmental and non-governmental hospitals, and no differences in accuracy between DNFs from hospitals using an electronic patient record system or those using paper-based documentation. The only statistically significant difference between governmental and non-governmental hospitals was in DNF-EDC agreement at ICD 4-character level (53% and 64% respectively), but this difference was not significant at ICD block level (data not shown). With the exception of EDC-PHIC comparison at a detailed level (11% using electronic records and 21% using paper records), the use of an electronic patient record system or traditional paper medical records by the hospital had no impact on agreement. The EDC-PHIC difference disappeared at block level.

Agreement by diagnostic group (Table 4)

DNF- EDC agreement

There were considerable differences in agreement between diagnostic groups. At both the ICD 4-character level and block level, the accuracy of the DNFs (DNF-EDC agreement) was best for external causes and neoplasms. The weakest agreement was for metabolic conditions. The differences between the West Bank and Gaza samples were not statistically significant.

DNF- PHIC agreement

At both ICD detailed level and ICD block level, agreement in the West Bank sample between the original DNFs and PHIC Registry data was best for cardiac diseases, but still weak. It was generally better in the Gaza sample, with the highest value for neoplasms and cerebrovascular conditions, both significantly better than in the West Bank sample. The lowest DNF-PHIC agreement in the West Bank sample was for anomalies and external causes, significantly lower than in the Gaza sample. In the Gaza sample, the lowest DNF-PHIC agreement was for perinatal causes. This difference between the samples was not statistically significant.

EDC-PHIC agreement

The accuracy of the underlying causes in the PHIC register, measured by EDC-PHIC agreement, was highest, but still moderate, for neoplasms and cardiac conditions in the West Bank sample. In the Gaza sample the highest agreement was for neoplasms (62%, significantly higher than in the West Bank sample). The lowest agreement was for metabolic conditions in the West Bank sample and for perinatal causes in the Gaza sample.

Table 4. Agreement in underlying cause by diagnostic group: ICD 4-character level, point estimates and 95% confidence intervals

	DNF-EDC				DNF-PHIC			
	West Bank		Gaza Strip		West Bank		Gaza Strip	
	PE (%)	95% CI	PE (%)	95% CI	PE (%)	95% CI	PE (%)	95% CI
Neoplasm	76	65-87	84	73-95	39	26-51	64	49-79
Metabolic	32	12-52	29	8-50	16	1-31	19	1-37
Cardiac	49	38-61	37	13-61	42	30-54	29	5-54
Cerebro-vascular	53	38-68	45	21-69	16	5-27	63	39-87
Perinatal	46	27-66	45	23-68	11	0-23	5	0-15
Anomaly	51	34-69	47	21-74	3	0-9	29	5-54
External	78	60-96	-		9	0-22	-	
Other	56	43-69	45	32-59	16	6-26	36	22-50

	EDC-PHIC			
	West Bank		Gaza Strip	
	PE (%)	95% CI	PE (%)	95% CI
Neoplasm	36	26-50	62	47-77
Metabolic	4	0-11	10	0-23
Cardiac	36	25-48	24	1-46
Cerebrovascular	7	0-15	26	5-48
Perinatal	11	0-21	5	0-15
Anomaly	3	0-9	24	1-46
External	9	0-22	-	
Other	10	2-18	31	18-45

DISCUSSION

In both the West Bank and Gaza, 30%-40% of DNFs had sequence-related errors combined with incorrect order or entries and several causes on the lowest completed line in Part 1, and more than one on some DNFs. From an international perspective, this is a high figure (5). Agreement between the underlying cause according to the EDC and the underlying cause of death actually registered by the PHIC was low: 19% (West Bank) and 31% (Gaza) at the most detailed level and 44% (both samples) even at ICD block level. EDC-PHIC agreement is lower than in other published studies on the accuracy of mortality statistics. For example, a validation of mortality

statistics in Cape Town showed accuracy of 55% at WHO tabulation list 1 level (103 groups, roughly corresponding to block level) (6). Other studies vary widely in their assessment of the accuracy of mortality statistics, (7) but since they generally refer to mortality statistics in high-resource countries and to specific causes of death, the results are difficult to compare to the present study.

As in several other studies on the quality of mortality statistics, accuracy varies between diagnostic groups (7-9). Accuracy in the causes of death in the PHIC register, indicated by EDC-PHIC agreement, was best, but still moderate, at 69% for cerebrovascular diseases in the West Bank sample. This is surprising since most studies show the highest agreement for neoplasms, which is also the case in the Gaza sample (83%) (9). In the West Bank sample, neoplasms show 54% of agreement, which is low in comparison with other studies. Both samples show low agreement for metabolic conditions, which is consistent with findings from other studies.

Although the present study adds to the accuracy of mortality statistics registries' literature, it has several limitations. First, the study was limited to hospital deaths. A future study is needed to examine the accuracy of DNFs for deaths outside the hospital context. Second, pairwise percent agreement used in data analysis does not take into account that agreement may occur by chance. Therefore, the percent agreement may be over reported. Third, we do not know the number of deaths that are not reported to the authorities at all. The presence of deaths that occurred several decades ago in the sample indicates that some deaths are unreported. It is important to estimate the number of non-reported deaths and determine if the causes of death in such cases differ significantly from those actually reported to the authorities.

CONCLUSIONS

Based on the study, the PHIC Cause of Death Registry gives a poor picture of the causes of death in both the West Bank and Gaza Strip. This hampers the assessment of important aspects of public health, such as the burden of disease, maternal mortality and the prevalence of congenital anomalies. The conclusion is that PHIC mortality statistics are not comparable to statistics from other countries that adhere to the ICD instructions more closely, and also that the medical precision of the statistics (measured by EDC-PHIC agreement) would be improved if the international coding and classification rules were adhered to. To address these points, the following measures are proposed:

1) Use of international coding software for coding and classification.

The use of internationally recognized coding software would bring several benefits. The software automatically applies ICD instructions for the selection and classification of causes of death. It also covers less common cases that coders may be unfamiliar with and brings coding and classification in line with the ICD instructions. A dictionary of medical terms is included in the software, contributing further to consistent coding. Further, coding software speeds up the coding process since it reduces the burden on the sole individual at PHIC who is currently responsible for all the coding. At present there can be a delay of several months from the time the DNF is received by the PHIC until it is coded and entered in the CoDR. A further advantage of automated coding is that with automated coding, the PHIC Registry would contain multiple causes of death for all deaths, and the potential for detailed analysis and monitoring would be considerably enhanced.

2) Training and supervision of cause of death certification at hospitals

Studies of the quality of mortality statistics often suggest that training for certifiers on how to complete the death certificate would resolve issues related to quality. An automated coding system would facilitate the training of new coders. Another proposed approach is that each death certificate should be written and signed by two physicians, or that a senior physician should review all death certificates issued at the hospital. In some locations, a specialist committee has been set up to review all hospitalizations that end with the death of the patient, and also to verify the accuracy of the death certificate.

3) Feedback from the statistical office to the individual hospital on problems listed in Table 2

It would be even more efficient if the hospital or certifier could be contacted in such cases and asked to amend the DNF.

4) Review and improved data handling at hospitals with electronic patient records

Electronic patient record systems are supposed to improve data availability and facilitate the administration of hospital care, but appeared to have the opposite result in our study. Legislation requires the DNF to be issued on paper; there is no digital DNF in the electronic system. Instead, paper DNFs are kept in a special file. As a result, several DNFs could not be retrieved from hospitals with electronic patient record systems. There were also other aberrations, for example deceased patients recorded as still being alive.

5) Electronic systems must be adapted to meet the legal requirements for reporting causes of death

6) The updating of basic variables such as whether the patient is alive or not should be facilitated.

In addition, steps should also be taken to improve the overall completeness of the Cause of Death Registry. According to a study conducted by the Palestinian central bureau of statistics, it now includes only about 60% of all deaths (3). Consequently, the statistics would still provide a poor picture of the causes of death even if the accuracy of reported cases was considerably improved.

Acknowledgements The authors thank M Abu Ghali, J Bitar, U Balawi, S Khammash, D Ghoushe, M Ali and E Aly.

Contributors LA conceptualised and designed the study. SM and IR were responsible for data collection, and LA and SM were responsible for analysis of the data. All the authors contributed to the preparation of the manuscript.

Competing interest: None declared

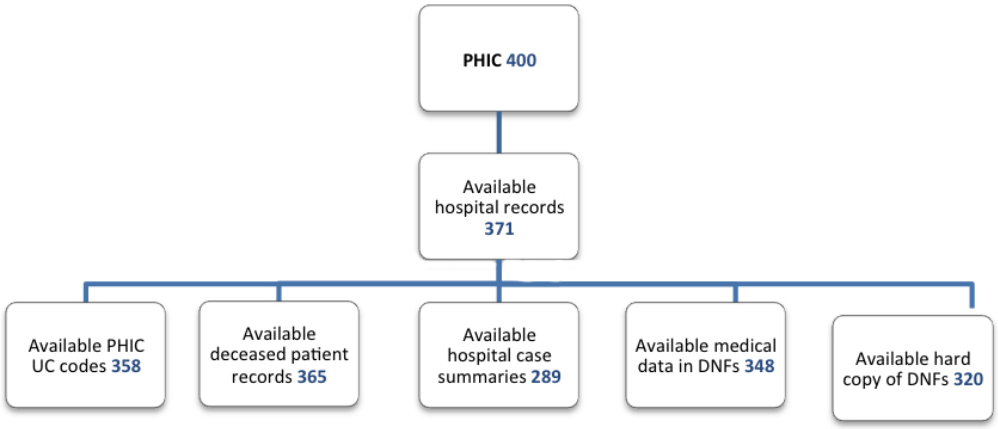
Ethical approval The Helsinki Committee for Ethical Approval- Palestinian Health Research Council, Gaza Strip/Palestine approved the study protocol in January 2014.

Data sharing statement No additional data available

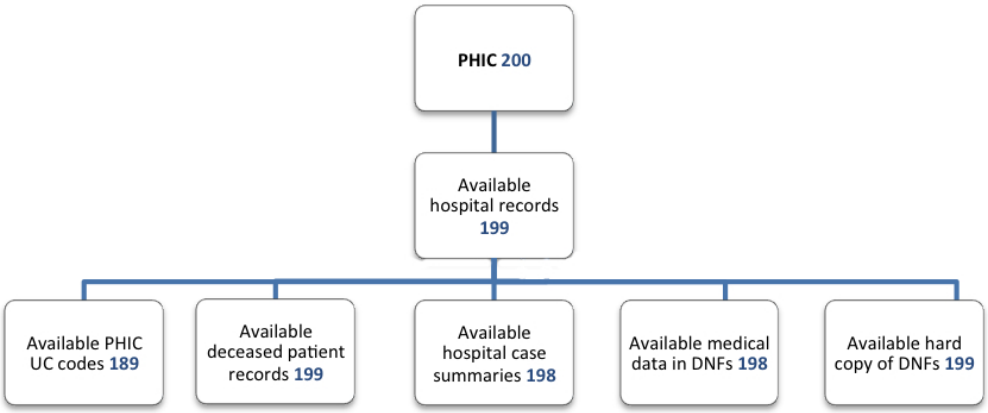
Funding This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

1. World Health Organization. Strengthening civil registration and vital statistics for births, deaths and causes of death: resource kit. 2013.
2. Sample XS. http://external.informer.com/brixtonhealth.com/samplexs.html
3. Palestinian Central Bureau of Statistics. Assessment of birth and mortality data in the West Bank 2011-2012. Ramallah: Palestinian Central Bureau of Statistics, 2013.
4. German Institute of Medical Documentation and Information. Iris Institute. Germany 2015 [cited 2016 October 18, 2016]; Available from: http://www.dimdi.de/static/en/klassi/irisinstitute/index.htm.
5. Jougla E, Rossolin F, Niyonsenga A, Chappert J, Johansson L, Pavillon G. Comparability and quality improvement in European causes of death statistics. Eurostat project. 2001;96:S99-5761.
6. Burger EH, Groenewald P, Bradshaw D, Ward AM, Yudkin PL, Volmink J. Validation study of cause of death statistics in Cape Town, South Africa, found poor agreement. Journal of clinical epidemiology. 2012;65(3):309-16.
7. Johansson LA, Westerling R, Rosenberg HM. Methodology of studies evaluating death certificate accuracy were flawed. Journal of clinical epidemiology. 2006;59(2):125-31.
8. Johansson LA, Björkenstam C, Westerling R. Unexplained differences between hospital and mortality data indicated mistakes in death certification: an investigation of 1,094 deaths in Sweden during 1995. Journal of clinical epidemiology. 2009;62(11):1202-9.
9. Johansson LA. Targeting non-obvious errors in death certificates. Doctoral dissertation, Acta Universitatis Upsaliensis. 2008.



141x61mm (150 x 150 DPI)



142x67mm (150 x 150 DPI)

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of *cross-sectional studies*

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	3
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	5-6
Objectives	3	State specific objectives, including any prespecified hypotheses	5
Methods			
Study design	4	Present key elements of study design early in the paper	6-8
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6-10
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants	6-10
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	7-8, 11-13
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	7-8, 11-13
Bias	9	Describe any efforts to address potential sources of bias	*13*
Study size	10	Explain how the study size was arrived at	6
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	11-13
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	11-13
		(*b*) Describe any methods used to examine subgroups and interactions	12-13
		(*c*) Explain how missing data were addressed	7-10
		(*d*) If applicable, describe analytical methods taking account of sampling strategy	11-13
		(*e*) Describe any sensitivity analyses	
Results			

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	7-8, 10
		(b) Give reasons for non-participation at each stage	NA
		(c) Consider use of a flow diagram	10
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	9
		(b) Indicate number of participants with missing data for each variable of interest	9-10
Outcome data	15*	Report numbers of outcome events or summary measures	14-19
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	14,16.19
		(*b*) Report category boundaries when continuous variables were categorized	16-19
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	17
Discussion			
Key results	18	Summarise key results with reference to study objectives	19-20
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	20
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	21-23
Generalisability	21	Discuss the generalisability (external validity) of the study results	20
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	23

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Accuracy of Mortality Statistics in Palestine: A retrospective cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-026640.R1
Article Type:	Research
Date Submitted by the Author:	28-Jan-2019
Complete List of Authors:	Massad, Salwa; Palestinian National Institute of Public Health, Dalloul, Hadil; Palestinian National Institute of public Health Ramlawi, Asad; Ministry of Health Rayyan, Izzat; Palestinian National Institute of Public Health Salman, Rand; Palestinian National Institute of Public Health Johansson, Lars; Uppsala University, Department of Public Health and Caring Sciences
Primary Subject Heading:	Public health
Secondary Subject Heading:	Health informatics
Keywords:	EPIDEMIOLOGY, Protocols & guidelines < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, PUBLIC HEALTH, cohort study

Title Page

Accuracy of Mortality Statistics in Palestine: A retrospective cohort study

Salwa G Massad1, Hadil Dalloul1, Asad Ramlawi2, Izzat Rayyan1, Rand Salman1, Lars Age Johansson3,*

[1]Palestinian National Institute of Public Health, Ramallah, Palestine, [2] Palestinian Ministry of Health, Ramallah, Palestine, [3]Department of Public Health and Caring Sciences, Uppsala University, Uppsala, Sweden.

*Correspondence to: Salwa G Massad, PhD. Research manager, Palestinian National Institute of Public Health, Hay El Razy Ramallah, West Bank/Palestine Tel: 00970598161071, Email: salwamassad@gmail.com

Email addresses:

Salwa G Massad (salwamassad@gmail.com)

Hadil Dalloul (hadil.dalloul@gmail.com)

Asad Ramlawi (ramlawi_asad@hotmail.com)

Izzat Rayyan (rayyan67@yahoo.com)

Rand Salman (salmanr@who.int)

Lars Age Johansson (lars.age@bredband.net)

Key words: cohort study, hospital based study, health system evaluation

Abbreviations:

CoDR: Cause of Death Registry

DNF: Death notification form

EDC: Medical Extraction Forms

1 ICD: International Classification of Diseases

2 MoH: Ministry of Health

3 MoI: Ministry of Interior

4 PHC: Primary Health Care Directorate

5 PHIC: Palestinian Health Information Centre

Accuracy of Mortality Statistics in Palestine: A retrospective cohort study

Abstract

Objective To examine the accuracy of mortality statistics in Palestine, to identify gaps, and to provide evidence based recommendations to improve mortality statistics in Palestine.

Study Design and Setting A retrospective death registry-based study that examined a stratified random sample of DNFs of patients who died in hospitals in Palestine and were reported in 2012. We randomly selected 600 deceased from the Cause of Death Registry: 400 from the West Bank and 200 from the Gaza Strip. Analysis was based on the randomly selected deaths that we were able to retrieve the medical records for ;371 deaths in the West Bank and 199 deaths in the Gaza Strip

Results Data in the Palestinian Health Information Center Registry (PHIC) had a low degree of accuracy: less than half of the underlying causes stated the correct cause of death. In general, deaths due to malignant neoplasms were more accurately reported on Death Notification Forms (DNFs) than other causes of death, and metabolic diseases (including diabetes) were the most problematic. Issues with coding and classification at the Palestinian Health Information Centre (PHIC) were most apparent for perinatal conditions and congenital anomalies.

Conclusion Procedures for coding and classification at the PHIC deviate considerably from the international norms defined in the International Classification of Diseases (ICD) and account to a considerable extent for the discrepancies between the cause of death determined on the medical data on the death extracted from the deceased patient's hospital records and the cause of death coded by the PHIC. We recommend the introduction of international coding software for coding

and classification, and a review to improve data handling in hospitals, especially those with electronic patient records.

Summary box

Strengths and limitations of this study

- The present study is the first assessment study of accuracy of mortality statistics in Palestine and to our knowledge; it is the first assessment study in the Middle East and North Africa Region.
- The present study is the first one to examine the completeness and accuracy of death notification forms.
- The study was limited to hospital deaths.
- The presence of deaths that occurred several decades ago in the sample indicates that reporting is incomplete. It is important to estimate the number of non-reported deaths and determine if the causes of death in such cases differ significantly from those actually reported to the authorities.

INTRODUCTION

Statistics on causes of death are essential indicators of the overall health and quality of life of a population. They can be used to monitor the status of health, estimate the burden of major diseases over time and across geographical regions, and to assess the impact of health interventions. However, this requires accurate statements on the cause of death and accurate statistics collected in accordance with international standards (1).

In Palestine, the Cause of Death Registry (CoDR) held by the Palestinian Health Information Centre (PHIC) at the Ministry of Health (MoH) is part of the national vital statistics database and holds information on deaths and the causes of deaths since 1994. Prior to that date, causes of death were registered with the Israeli Civil Administration. ICD has been used for coding since 1960s and in early 1999, PHIC started using ICD10 instead of ICD9.

The main source of data for the CoDR is the Death Notification Form (DNF), issued by the attending physician. The DNF includes information about the main and underlying causes of death. The DNF is given to the family of the deceased, who are the ones responsible for notifying the death to the Primary Health Care Directorate (PHC) in each district. In the Gaza Strip, some hospitals keep a copy of the DNF but others do not. The PHC or the family of the deceased sends the DNF to the Ministry of Interior (MoI). At the end of each month, the PHC Directorate compiles all DNFs and mails them to the PHIC for coding and registration. The PHIC also receives some DNFs from governmental hospitals (Figure 1). For neonatal deaths, there are no DNFs; instead, a monthly report of neonatal deaths is sent by each hospital to the PHIC. The PHIC codes the underlying causes of death and enters the death into the Cause of Death Registry (CoDR).

Figure 1: Palestinian Cause of Death Statistical System

We carried out this study to examine the accuracy of mortality statistics in Palestine. We examined hospital deaths, which accounts for 50% -60% of deaths in Palestine. This study will serve to identify information gaps, and to provide evidence based recommendations to improve mortality statistics in Palestine.

METHODOLOGY

Study design

This retrospective death registry-based study examined a stratified random sample of DNFs of patients who died in hospitals in the West Bank and Gaza Strip and were reported in 2012. There are 81 hospitals in total in Palestine, with 51 in the West Bank (7 of them are in Jerusalem) and 30 in the Gaza Strip, serving around 4.78 million Palestinian (2.88 million in the West Bank and 1.9 million in Gaza Strip). Around 33% of hospitals in Palestine are governmental. The Ministry of Health accounts for 44% of the bed capacity in the West Bank, and 69% of the bed capacity in the Gaza Strip. We randomly selected 600 deceased from the Cause of Death Registry: 400 from the West Bank and 200 from the Gaza Strip. The sample size was estimated using the SampleXS program (2) and based on the following assumptions: the number of registered deaths in 2012 was 12,000; and the completeness of death certificates was 60% based

on the recent PCBS study of the completeness of the Cause of Death Registry (3); the margin of error 5%; and the design effect 1.5.

About 70% of the selected deaths in the West Bank occurred in government hospitals, while in the Gaza Strip all deaths were in government hospitals. We took a random sample of hospital deaths from the following four age groups: less than 1 year, 1-17 years, 18-64 years and 65+ years from Cause of Death Registry at PHIC. The study was carried out in 2014 in close collaboration with the Ministry of Health (MoH) in the West Bank and Gaza Strip. The response rate was 100%; all selected hospitals in the West Bank and Gaza Strip agreed to participate in the study.

Material

We used several sources of data to assess the accuracy and completeness of the mortality data (Figure 2). For each death in the sample we collected the following data:

(i) The Death Notification Form (DNF)

- Scanned image of the DNF
- Formal assessment of the DNF, such as legibility of the handwriting and if the administrative data had been filled out correctly with the certifier's signature.
- The causes of death in free text and administrative information were extracted from the DNF.

(ii) Data from PHIC:

- Date on which the death was registered.

- International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD 10) codes for underlying causes of death as recorded by the PHIC in the Death Registry.

(iii) Medical Extraction Forms (EDC):

Medical data on the death extracted from the deceased patient's hospital records. A team of specially trained physicians from each hospital in the study extracted data from the hospital record on the train of events leading to the death. The extraction file contained data on the reason for hospitalization, previous medical history and the course of events during hospitalization.

(iv) Hospital case summaries:

- Scanned image of the original hospital case summary.
- Case summary based on hospital records, completed by specially trained physicians in each hospital.

Figures 3 and 4 summarize data availability and attrition in the West Bank and Gaza Strip. In the West Bank, hospital records could be found for 371 cases of the 400 randomly selected deaths. In the Gaza Strip, we found hospital records for 199 of the 200 randomly selected deaths. We included all deaths reported in 2012 irrespective of date of death. The PHIC Registry gave an ICD code for the underlying cause of death in 358 of the 371 cases in the West Bank sample for which hospital records were available. In the Gaza Strip, 189 of the 199 cases in the sample had an ICD code for the underlying cause of death and 70 cases also included multiple causes of death registered by the PHIC. No multiple causes were present in the West Bank sample. Scanned images of the DNFs were available for 320 deaths in the West Bank

sample. EDC data on file were available for 365 West Bank deaths and for all Gaza Strip deaths. Table 1 summarizes the characteristics of the data.

Figure 2: Study Design Flow Chart

Table 1. Data characteristics of the records available for the study

Characteristic	N (%)	
	West Bank	**Gaza**
Sample size	**371**	**199**
- Government hospitals	265 (71%)	199(100%)
- Hospitals using electronic patient records	85 (23%)	0 (0%)
- Age		
< 1 year	75 (20%)	54 (27%)
1-17 years	45 (12%)	19 (10%)
18-64 years	149 (40%)	72 (36%)
65+	102 (27%)	54 (27%)
- Gender		
Male	200 (54%)	99 (50%)
Female	171 (46%)	100 (50%)

Figure 3: Data availability and attrition in West Bank

Figure 4: Data availability and attrition in Gaza Strip

STATISTICAL ANALYSIS

Completeness and timeliness

We did all analysis by region, due to the legislative and physical division of the occupied Palestinian territory, in terms of the separation of the Gaza Strip from the West Bank; resulting in two separate entities with different health systems and registries. We examined the proportion of complete DNFs in terms of personal data, immediate and underlying causes of death, and the attending physician and notifying person. We also examined the timeliness of registration, measured from the time of death to the time the deceased was registered in the Cause of Death Registry.

Underlying cause of death based on the patient's hospital file and on the DNF

Using the EDC, we determined the underlying cause of death and coded it using ICD-10. All coding, both of the original DNFs and of the EDC data, was done using Iris coding software (4). The version used (V4) corresponds to the 2013 edition of ICD-10. Iris fetches an ICD code for each diagnostic expression from a dictionary. Next, Iris prepares an input string of ICD codes for the ACME module (developed by the US National Center for Health Statistics). Finally, ACME selects an underlying cause according to the ICD rules and guidelines. After review and coding, we compared the three underlying causes – the one coded by PHIC; the one stated on the DNF form, but coded and classified by Iris; and the one according to the EDC and coded by researcher using IRIS– pair by pair:

- *Underlying cause on the DNF, coded by Iris vs Underlying cause in the PHIC Registry*

This comparison measures PHIC compliance with international rules for selection of the underlying cause of death.

- *Underlying cause on the DNF, coded by Iris vs Underlying cause according to the EDC*

This measures the accuracy of the original DNF.

- *Underlying cause according to the EDC vs Underlying cause in the PHIC Registry*

This comparison measures the overall accuracy of the underlying cause registered in the PHIC Registry.

We compared the underlying causes at two levels of detail: ICD detailed level (henceforth referred to as ICD 4-character level) and ICD block level. The principal use of measurements at detailed level data is to estimate the precision of the coding and classification, while measurements at ICD block level provide a broad assessment of the general trustworthiness for public health purposes.

For cross tabulation, we examined the accuracy of mortality statistics by gender and age group of the deceased patient, hospital affiliation (governmental and non-governmental), type of hospital records and underlying cause of death using chi square test. Tests of significance were two-sided with p- value ≤ 0.05. We divided the underlying causes according to the EDC into eight main diagnostic groups based on the frequency of the most common underlying causes according to the EDCs.

- Neoplasms (ICD-10 Chapter II)
- Metabolic diseases (ICD-10 Chapter IV)
- Cardiac diseases (ICD-10 I00-I51)
- Cerebrovascular diseases (ICD-10 I60-I69)
- Perinatal conditions (ICD-10 Chapter XVI)
- Congenital anomalies (ICD-10 Chapter XVII)

- External causes (ICD-10 Chapter XX)
- Other causes of death (ICD-10 chapters and codes not included elsewhere).

("Other" includes infectious diseases, diseases of the blood, neurological diseases, diseases affecting vision and hearing, respiratory diseases, gastrointestinal diseases, skin diseases, musculoskeletal diseases, urogenital diseases and symptoms with no clear connection to a single underlying disease.)

Selecting the underlying cause of death

The classification expert (LAJ) checked the diagnoses registered in the patients' files against the DNFs and corrected the registered text if it differed from theDNF. Similarly, LAJ checked the causes of death reported on the EDCs against the case summaries and against the scanned images of the hospital summaries. If an EDC was inconsistent with the case summary or the hospital summary, the EDC was corrected according to the description in the summaries.

Also, for each DNF, the following markers were used to record non-compliance with WHO instructions for completion of the death certificate: abbreviations used in the medical section, illegible writing in the medical section, sequencing errors in Part 1, symptomatic or secondary condition reported as underlying cause of death, cause of death insufficiently specified, and other reporting errors (such as reporting several causes on the lowest completed line in Part 1, placing the underlying cause in Part 2 of the certificate, or using the wrong field on the DNF to report causes of death).

Patient and Public Involvement

Not applicable for this study as we used mortality statistics and data was de-identified prior to data analysis.

RESULTS

Completeness of DNF forms and other characteristics

Table 2 gives a summary of other characteristics of DNFs unrelated to the accuracy of the underlying cause of death. Although most deaths had been reported within a year (median of 89 days; range excluding extreme values 0-365 days), occasionally deaths from the 1980s and 1990s were only reported to the authorities in 2012. A fairly high number of DNFs (23% West Bank, 26% Gaza) reported an underlying cause of death that presumably developed as a complication of some other condition, yet the underlying condition is not recorded. For example, in several DNFs, kidney failure was reported as the underlying cause of death, but there was no mention of the reason why kidney failure developed, such as diabetes, glomerulonephritis or urinary obstruction. Administrative data (Part 1 of the DNF) were complete in only 7% and 2% of DNFs from hospitals with electronic patient records and paper patient records respectively. Other problems with electronic records included several deaths that had been registered twice or more, and some of the deceased recorded in the hospital files were still alive.

Table 2. Completeness and other DNF characteristics

	West Bank (N=320)		Gaza Strip (N=199)	
	cases	%	cases	%
No medical data on DNF	23	7.2	1	0.5
- electronic records (71)	15	21.1	-	-
- paper records (249)	8	3.2	1	0.5
Administrative data complete (Part I)	9	2.8	108	54.3
- electronic records (71)	5	7.0	-	-
- paper records (249)	4	1.6	-	-
Abbreviations used	132	41.2	76	38.2
Illegible writing	96	30.0	11	5.5
Sequence errors	64	20.0	38	19.1
Non-informative UC	74	23.1	51	25.6
UC lacking specificity	36	11.2	12	6.0
Irrelevant information	105	32.8	22	11.0
Other certification errors	92	28.8	40	20.1

Competing causes of death	36	11.2	23	11.6
Certifier's signature complete	159	59.7	50	25.1
Registry delay (days; median)	89		57	

Agreement between the three underlying causes (Table 3)

DNF- EDC agreement

Agreement on the underlying cause was highest between the underlying cause derived from the original DNF and the medical extraction forms (DNF-EDC). At the most detailed level, the agreement was 56% for the West Bank sample and 52% for the Gaza Strip sample. The difference between the two sample groups was not statistically significant (p-value: 0.37).

DNF-PHIC agreement

DNF-PHIC agreement between the DNF coded in line with international standards and the underlying cause as coded by the PHIC was 23% in the West Bank sample at ICD 4-character level and 39% in the Gaza sample. The difference between the samples was statistically significant (p-value: 0.0001). At ICD block level, the West Bank sample had a DNF-PHIC agreement of 46% and the Gaza sample 53%, a difference that was not statistically significant (p-value:0.12). .

EDC-PHIC agreement

Agreement was lowest between the EDC and the underlying cause recorded in the PHIC Registry. At ICD 4-character level, EDC-PHIC agreement was 19% in the West Bank sample and 31% in the Gaza sample.

Table 3. Agreement on underlying cause by ICD level: point estimates, and for ICD-4 character level and ICD block level and 95% confidence intervals

		ICD 4-character		ICD block level	
		PE (%)	95% CI	PE (%)	95% CI
DNF-EDC	West Bank	56	51-62	68	63-73
	Gaza Strip	52	44-59	62	55-68
DNF-PHIC	West Bank	23	19-28	46	41-51
	Gaza Strip	39	31-46	53	46-60
EDC-PHIC	West Bank	19	15-23	44	38-49
	Gaza Strip	31	24-38	44	37-51

Agreement by sex, age, type of hospital and type of deceased hospital records

In both samples, there was no noticeable relationship between the sex and age of the decedent and the accuracy of the cause of death There were no statistically significant differences between males and females in underlying cause agreement (21% vs.16%, respectively, p-value: 0.16). Also, there were few significant differences between the four age groups, either at ICD 4-character or ICD block level. At a detailed level there was, significantly lower agreement for the youngest age groups in the two comparisons involving the codes in the PHIC Registry (DNF-PHIC and EDC-PHIC) (7% and 8%) compared to 27% and 24% at age group 18-64years (both p-value: <0.0001). At block level, only the DNF-PHIC comparison showed a significant difference between children <1 year and older people >64 (43% vs 60% respectively, p-value: 0.0002), probably reflecting difficulties in the coding of congenital anomalies and perinatal causes.

Also, the West Bank sample showed only minor differences between governmental and non-governmental hospitals, and no differences in accuracy between DNFs from hospitals using an electronic patient record system or those using paper-based documentation. The only statistically significant difference between governmental and non-governmental hospitals was in DNF-EDC agreement at ICD 4-character level (53% and 64% respectively, p-value: 0.013), but this difference was not significant at ICD block level (data not shown). With the exception of EDC-PHIC comparison at a detailed level (11% using electronic records and 21% using paper records, p-value: 0.0003), the use of an electronic patient record system or traditional paper medical records by the hospital had no impact on agreement. The EDC-PHIC difference disappeared at block level.

Agreement by diagnostic group (Table 4)

DNF- EDC agreement

There were considerable differences in agreement between diagnostic groups. At both the ICD 4-character level and block level, the accuracy of the DNFs (DNF-EDC agreement) was best for external causes and neoplasms. The weakest agreement was for metabolic conditions. The differences between the West Bank and Gaza samples were not statistically significant (56%, 52% respectively, p-value:0.37).

DNF- PHIC agreement

At both ICD detailed level and ICD block level, agreement in the West Bank sample between the original DNFs and PHIC Registry data was best for cardiac diseases, but still weak. It was generally better in the Gaza sample, with the highest value for neoplasms (39% , 64% respectively, p-value: <0.0001) and cerebrovascular conditions (16% , 63% respectively, p-value: <0.0001), both significantly better than in the West Bank sample. The lowest DNF-PHIC agreement in the West Bank sample was for anomalies and external causes significantly lower than in the Gaza sample (3%, 29% respectively, p-value:<0.0001). In the Gaza sample, the lowest DNF-PHIC agreement was for perinatal causes.

EDC-PHIC agreement

The accuracy of the underlying causes in the PHIC register, measured by EDC-PHIC agreement, was highest, but still moderate, for neoplasms and cardiac conditions in the West Bank sample. In the Gaza sample the highest agreement was for neoplasms (62%, significantly higher than in

the West Bank sample 36%, p-value :< 0.0001). The lowest agreement was for metabolic conditions in the West Bank sample and for perinatal causes in the Gaza sample.

Table 4. Agreement in underlying cause by diagnostic group: ICD 4-character level, point estimates and 95% confidence intervals

	DNF-EDC				DNF-PHIC			
	West Bank		Gaza Strip		West Bank		Gaza Strip	
	PE (%)	95% CI	PE (%)	95% CI	PE (%)	95% CI	PE (%)	95% CI
Neoplasm	76	65-87	84	73-95	39	26-51	64	49-79
Metabolic	32	12-52	29	8-50	16	1-31	19	1-37
Cardiac	49	38-61	37	13-61	42	30-54	29	5-54
Cerebro-vascular	53	38-68	45	21-69	16	5-27	63	39-87
Perinatal	46	27-66	45	23-68	11	0-23	5	0-15
Anomaly	51	34-69	47	21-74	3	0-9	29	5-54
External	78	60-96	-		9	0-22	-	
Other	56	43-69	45	32-59	16	6-26	36	22-50

	EDC-PHIC			
	West Bank		Gaza Strip	
	PE (%)	95% CI	PE (%)	95% CI
Neoplasm	36	26-50	62	47-77
Metabolic	4	0-11	10	0-23
Cardiac	36	25-48	24	1-46
Cerebrovascular	7	0-15	26	5-48
Perinatal	11	0-21	5	0-15
Anomaly	3	0-9	24	1-46
External	9	0-22	-	
Other	10	2-18	31	18-45

DISCUSSION

This is the first baseline study of accuracy of mortality statistics in Palestine. Based on the study findings, the accuracy is low, mostly due to inaccuracy in coding the underlying cause of death. Add to that, in both the West Bank and Gaza, 30%-40% of DNFs had sequence-related errors combined with incorrect order or entries and competing causes of death. From an international perspective, this is a high figure (5). Agreement between the underlying cause according to the EDC and the underlying cause of death actually registered by the PHIC was low: 19% (West Bank) and 31% (Gaza) at the most detailed level and 44% (both samples) even at ICD block

level. In some comparisons the Gaza strip performed better than the West Bank, in others it was the other way round. Further research is needed to understand the differences in accuracy between Gaza Strip and West Bank

Based on study findings, the accuracy of mortality statistics is lower than in other published similar studies. For example, a validation of mortality statistics in Cape Town showed accuracy of 55% at WHO tabulation list 1 level (103 groups, roughly corresponding to block level) (6). Other studies vary widely in their assessment of the accuracy of mortality statistics, (7) but since they generally refer to mortality statistics in high-resource countries and to specific causes of death, the results are difficult to compare to the present study.

As in several other studies on the quality of mortality statistics, accuracy varies between diagnostic groups (7-9). Accuracy in the causes of death in the PHIC register, indicated by EDC-PHIC agreement, was best, but still moderate, at 69% for cerebrovascular diseases in the West Bank sample. This is surprising since most studies show the highest agreement for neoplasms, which is also the case in the Gaza sample (83%) (9). In the West Bank sample, neoplasms show 54% of agreement, which is low in comparison with other studies. Both samples show low agreement for metabolic conditions, which is consistent with findings from other studies.

Although the present study adds to the accuracy of mortality statistics registries' literature, it has several limitations. First, the study was limited to hospital deaths. A future study is needed to examine the accuracy of DNFs for deaths outside the hospital context. Second, pairwise percent agreement used in data analysis does not take into account that agreement may occur by chance. Therefore, the percent agreement may be over reported. Third, we do not know the number of deaths that are not reported to the authorities at all. The presence

of deaths that occurred several decades ago in the sample indicates that some deaths are unreported. It is important to estimate the number of non-reported deaths and determine if the causes of death in such cases differ significantly from those actually reported to the authorities.

CONCLUSIONS

Based on the study, the PHIC Cause of Death Registry gives a poor picture of the causes of death in both the West Bank and Gaza Strip. This hampers the assessment of important aspects of public health, such as the burden of disease, maternal mortality and the prevalence of congenital anomalies. The conclusion is that PHIC mortality statistics are not comparable to statistics from other countries that adhere to the ICD instructions more closely, and also that the medical precision of the statistics (measured by EDC-PHIC agreement) would be improved if the international coding and classification rules were adhered to. To address these points, the following measures are proposed:

1) Use of international coding software for coding and classification (10).

The use of internationally recognized coding software would bring several benefits. The software automatically applies ICD instructions for the selection and classification of causes of death. It also covers less common cases that coders may be unfamiliar with and brings coding and classification in line with the ICD instructions. A dictionary of medical terms is included in the software, contributing further to consistent coding. Further, coding software speeds up the coding process since it reduces the burden on the limited staff at PHIC who are currently responsible for all the coding. At present there can be a delay of several months from the time the DNF is received by the PHIC until it is coded and entered in the CoDR. A further advantage of automated coding is that with automated coding, the PHIC Registry would contain multiple

causes of death for all deaths, and the potential for detailed analysis and monitoring would be considerably enhanced.

2) Training and supervision of cause of death certification at hospitals

Studies of the quality of mortality statistics often suggest that training for certifiers on how to complete the death certificate would resolve issues related to quality. An automated coding system would facilitate the training of new coders. Another proposed approach is that each death certificate should be written and signed by two physicians (11, 12), or that a senior physician should review all death certificates issued at the hospital. In some locations, a specialist committee has been set up to review all hospitalizations that end with the death of the patient, and also to verify the accuracy of the death certificate (12, 13).

3) Feedback from PHIC to the individual hospital on problems listed in Table 2 (14,15). It would be even more efficient if the hospital or certifier could be contacted in such cases and asked to amend the DNF.

In addition, steps should also be taken to improve the overall completeness of the Cause of Death Registry. First, review and improve data handling at hospitals with electronic patient records. Electronic patient record systems are supposed to improve data availability and facilitate the administration of hospital care, but appeared to have the opposite result in our study. Second, develop a digital DNF in the electronic system. Second, enforce completion of the death report in the electronic system by adding specific restrictions in the electronic system.

According to a study conducted by the Palestinian Central Bureau of Statistics, it now includes only about 60% of all deaths (3). Consequently, the statistics would still provide a poor picture of the causes of death even if the accuracy of reported cases was considerably improved.

Acknowledgements The authors thank M Abu Ghali, J Bitar, U Balawi, S Khammash, D Ghoushe, M Ali and E Aly.

Contributors LA conceptualised and designed the study. SM and IR were responsible for data collection. LA, SM and HD were responsible for analysis of the data. SM, LA, HD, IR, RS, AR contributed to the preparation of the manuscript, revised, and approved the final draft.

Competing interest: None declared

Ethical approval The Helsinki Committee for Ethical Approval- Palestinian Health Research Council, Gaza Strip/Palestine approved the study protocol in January 2014 (Ethical approval number: PHRC/HC/67/14).

Data sharing statement No additional data available.

Funding This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

1. World Health Organization. Strengthening civil registration and vital statistics for births, deaths and causes of death: resource kit. 2013.
2. Sample XS. http://external.informer.com/brixtonhealth.com/samplexs.html
3. Palestinian Central Bureau of Statistics. Assessment of birth and mortality data in the West Bank 2011-2012. Ramallah: Palestinian Central Bureau of Statistics, 2013.
4. German Institute of Medical Documentation and Information. Iris Institute. Germany 2015 [cited 2016 October 18, 2016]; Available from: http://www.dimdi.de/static/en/klassi/irisinstitute/index.htm.
5. Jougla E, Rossolin F, Niyonsenga A, Chappert J, Johansson L, Pavillon G. Comparability and quality improvement in European causes of death statistics. Eurostat project. 2001;96:S99-5761.
6. Burger EH, Groenewald P, Bradshaw D, Ward AM, Yudkin PL, Volmink J. Validation study of cause of death statistics in Cape Town, South Africa, found poor agreement. Journal of clinical epidemiology. 2012;65(3):309-16.
7. Johansson LA, Westerling R, Rosenberg HM. Methodology of studies evaluating death certificate accuracy were flawed. Journal of clinical epidemiology. 2006;59(2):125-31.
8. Johansson LA, Björkenstam C, Westerling R. Unexplained differences between hospital and mortality data indicated mistakes in death certification: an investigation of 1,094 deaths in Sweden during 1995. Journal of clinical epidemiology. 2009;62(11):1202-9.
9. Johansson LA. Targeting non-obvious errors in death certificates. Doctoral dissertation, Acta Universitatis Upsaliensis. 2008.

10. German Institute of Medical Documentation and Information. Iris Institute. Secondary Iris Institute 2018. https://www.dimdi.de/dynamic/en/classifications/iris-institute/.

11. Lahti RA, Penttilä A. Cause-of-death query in validation of death certification by expert panel; effects on mortality statistics in Finland, 1995. Forensic science international 2003;131(2-3):113-24.

12. Hanzlick R. Quality assurance review of death certificates: a pilot study. The American journal of forensic medicine and pathology 2005;26(1):63-65.

13. Curb JD, Davis BR, Tung B. Standardized physician preparation of death certificates. Controlled Clinical Trials 1985;6(3):237.

14. Jougla E. Comparability and Quality Improvement of European Causes of Death Statistics. European Commission DG Sanco Agreement. Final report July 2001. EDC DGV/F3 SOC 98 20108.

15. Hoyert DL, Lima AR. Querying of death certificates in the United States. Public Health Rep. 2005;120(3):288-93.

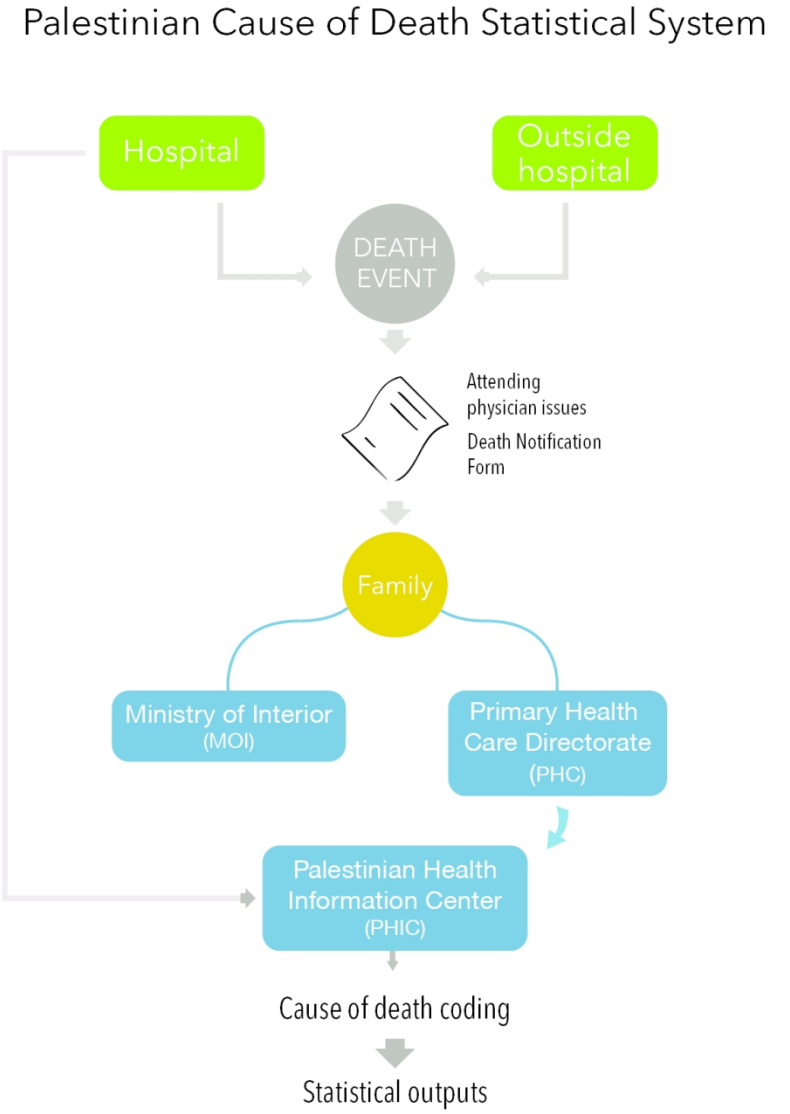


Figure 1: Palestinian Cause of Death Statistical System

209x297mm (150 x 150 DPI)

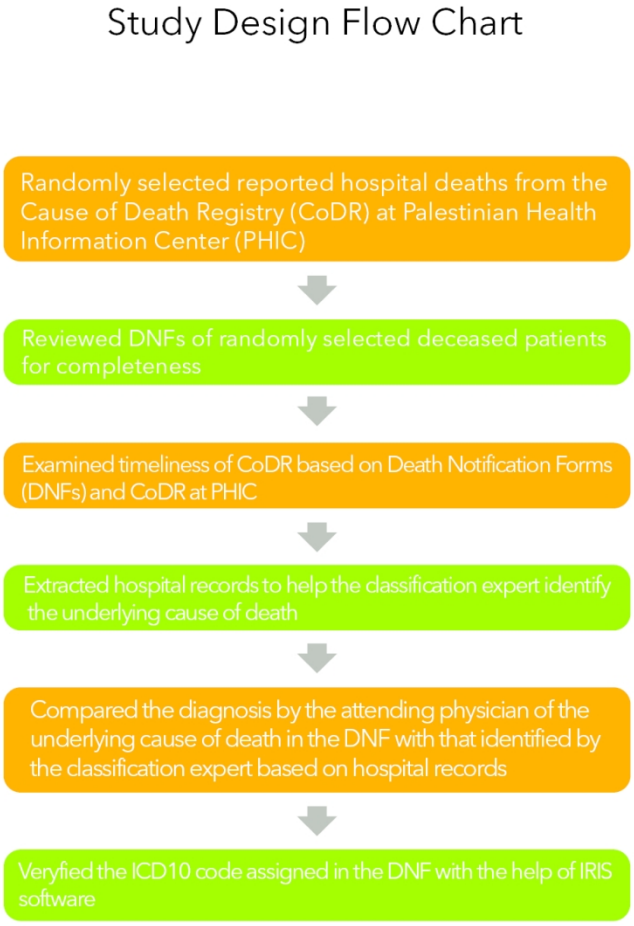


Figure 2: Study Design Flow Chart

209x297mm (150 x 150 DPI)

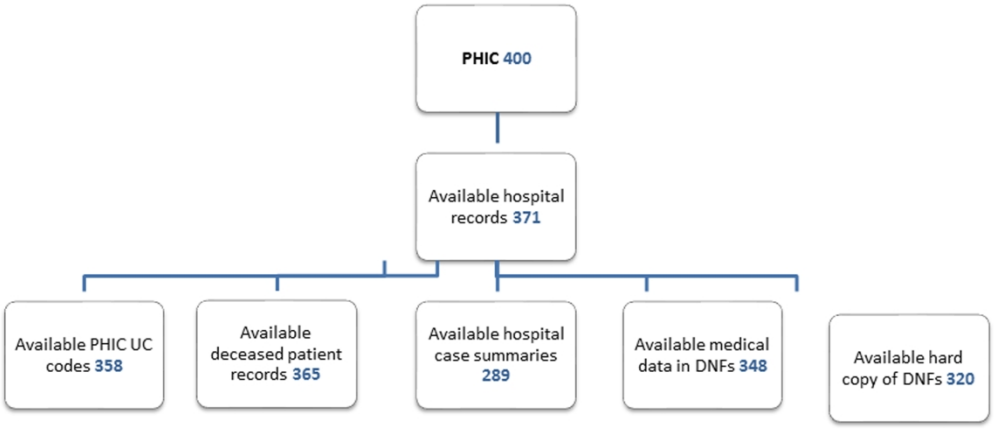


Figure 3: Data availability and attrition in West Bank

297x155mm (300 x 300 DPI)

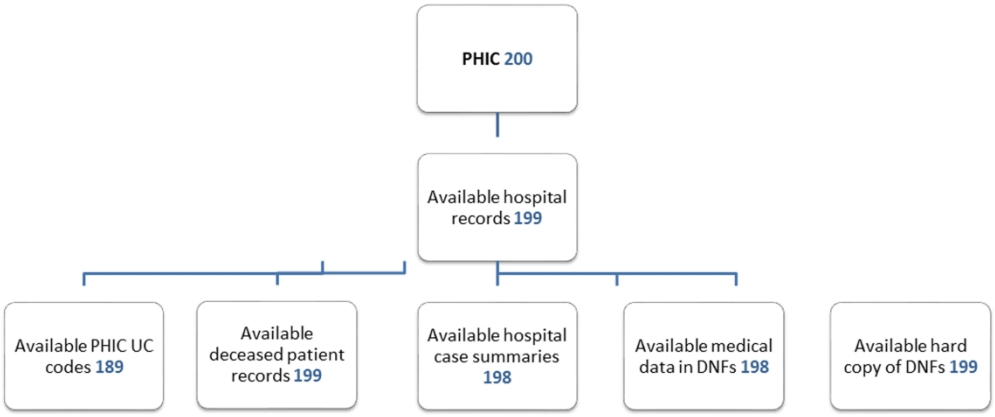


Figure 4: Data availability and attrition in Gaza Strip

297x155mm (300 x 300 DPI)

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of *cross-sectional studies*

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	3
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	5-6
Objectives	3	State specific objectives, including any prespecified hypotheses	5
Methods			
Study design	4	Present key elements of study design early in the paper	6-8
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6-10
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants	6-10
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	7-8, 11-13
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	7-8, 11-13
Bias	9	Describe any efforts to address potential sources of bias	*13*
Study size	10	Explain how the study size was arrived at	6
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	11-13
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	11-13
		(*b*) Describe any methods used to examine subgroups and interactions	12-13
		(*c*) Explain how missing data were addressed	7-10
		(*d*) If applicable, describe analytical methods taking account of sampling strategy	11-13
		(*e*) Describe any sensitivity analyses	
Results			

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	7-8, 10
		(b) Give reasons for non-participation at each stage	NA
		(c) Consider use of a flow diagram	10
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	9
		(b) Indicate number of participants with missing data for each variable of interest	9-10
Outcome data	15*	Report numbers of outcome events or summary measures	14-19
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	14,16.19
		(*b*) Report category boundaries when continuous variables were categorized	16-19
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	17
Discussion			
Key results	18	Summarise key results with reference to study objectives	19-20
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	20
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	21-23
Generalisability	21	Discuss the generalisability (external validity) of the study results	20
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	23

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Accuracy of Mortality Statistics in Palestine: A retrospective cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-026640.R2
Article Type:	Research
Date Submitted by the Author:	07-Feb-2019
Complete List of Authors:	Massad, Salwa; Palestinian National Institute of Public Health, Dalloul, Hadil; Palestinian National Institute of public Health Ramlawi, Asad; Ministry of Health Rayyan, Izzat; Palestinian National Institute of Public Health Salman, Rand; Palestinian National Institute of Public Health Johansson, Lars; Uppsala University, Department of Public Health and Caring Sciences
Primary Subject Heading:	Public health
Secondary Subject Heading:	Health informatics
Keywords:	EPIDEMIOLOGY, Protocols & guidelines < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, PUBLIC HEALTH, cohort study

Title Page

Accuracy of Mortality Statistics in Palestine: A retrospective cohort study

Salwa G Massad1, Hadil Dalloul1, Asad Ramlawi2, Izzat Rayyan1, Rand Salman1, Lars Age Johansson3,*

[1]Palestinian National Institute of Public Health, Ramallah, Palestine, [2] Palestinian Ministry of Health, Ramallah, Palestine, [3]Department of Public Health and Caring Sciences, Uppsala University, Uppsala, Sweden.

*Correspondence to: Salwa G Massad, PhD. Research manager, Palestinian National Institute of Public Health, Hay El Razy Ramallah, West Bank/Palestine Tel: 00970598161071, Email: salwamassad@gmail.com

Email addresses:

Salwa G Massad (salwamassad@gmail.com)

Hadil Dalloul (hadil.dalloul@gmail.com)

Asad Ramlawi (ramlawi_asad@hotmail.com)

Izzat Rayyan (rayyan67@yahoo.com)

Rand Salman (salmanr@who.int)

Lars Age Johansson (lars.age@bredband.net)

Key words: cohort study, hospital based study, health system evaluation

Abbreviations:

CoDR: Cause of Death Registry

DNF: Death notification form

EDC: Medical Extraction Forms

1 ICD: International Classification of Diseases

2 MoH: Ministry of Health

3 MoI: Ministry of Interior

4 PHC: Primary Health Care Directorate

5 PHIC: Palestinian Health Information Centre

Accuracy of Mortality Statistics in Palestine: A retrospective cohort study

Abstract

Objective To examine the accuracy of mortality statistics in Palestine, to identify gaps, and to provide evidence based recommendations to improve mortality statistics in Palestine.

Study Design and Setting A retrospective death registry-based study that examined a stratified random sample of DNFs of patients who died in hospitals in Palestine and were reported in 2012. We randomly selected 600 deceased from the Cause of Death Registry: 400 from the West Bank and 200 from the Gaza Strip. Analysis was based on the randomly selected deaths that we were able to retrieve the medical records for ;371 deaths in the West Bank and 199 deaths in the Gaza Strip

Results Data in the Palestinian Health Information Center Registry (PHIC) had a low degree of accuracy: less than half of the underlying causes stated the correct cause of death. In general, deaths due to malignant neoplasms were more accurately reported on Death Notification Forms (DNFs) than other causes of death, and metabolic diseases (including diabetes) were the most problematic. Issues with coding and classification at the Palestinian Health Information Centre (PHIC) were most apparent for perinatal conditions and congenital anomalies.

Conclusion Procedures for coding and classification at the PHIC deviate considerably from the international norms defined in the International Classification of Diseases (ICD) and account to a considerable extent for the discrepancies between the cause of death determined on the medical data on the death extracted from the deceased patient's hospital records and the cause of death coded by the PHIC. We recommend the introduction of international coding software for coding

and classification, and a review to improve data handling in hospitals, especially those with electronic patient records.

Summary box

Strengths and limitations of this study

- The present study is the first assessment study of accuracy of mortality statistics in Palestine and to our knowledge; it is the first assessment study in the Middle East and North Africa Region.
- The present study is the first one to examine the completeness and accuracy of death notification forms.
- The study was limited to hospital deaths.
- The presence of deaths that occurred several decades ago in the sample indicates that reporting is incomplete. It is important to estimate the number of non-reported deaths and determine if the causes of death in such cases differ significantly from those actually reported to the authorities.

INTRODUCTION

Statistics on causes of death are essential indicators of the overall health and quality of life of a population. They can be used to monitor the status of health, estimate the burden of major diseases over time and across geographical regions, and to assess the impact of health interventions. However, this requires accurate statements on the cause of death and accurate statistics collected in accordance with international standards (1).

In Palestine, the Cause of Death Registry (CoDR) held by the Palestinian Health Information Centre (PHIC) at the Ministry of Health (MoH) is part of the national vital statistics database and holds information on deaths and the causes of deaths since 1994. Prior to that date, causes of death were registered with the Israeli Civil Administration. ICD has been used for coding since 1960s and in early 1999, PHIC started using ICD10 instead of ICD9.

The main source of data for the CoDR is the Death Notification Form (DNF), issued by the attending physician. The DNF includes information about the main and underlying causes of death. The DNF is given to the family of the deceased, who are the ones responsible for notifying the death to the Primary Health Care Directorate (PHC) in each district. In the Gaza Strip, some hospitals keep a copy of the DNF but others do not. The PHC or the family of the deceased sends the DNF to the Ministry of Interior (MoI). At the end of each month, the PHC Directorate compiles all DNFs and mails them to the PHIC for coding and registration. The PHIC also receives some DNFs from governmental hospitals (Figure 1). For neonatal deaths, there are no DNFs; instead, a monthly report of neonatal deaths is sent by each hospital to the PHIC. The PHIC codes the underlying causes of death and enters the death into the Cause of Death Registry (CoDR).

Figure 1: Palestinian Cause of Death Statistical System

We carried out this study to examine the accuracy of mortality statistics in Palestine. We examined hospital deaths, which accounts for 50% -60% of deaths in Palestine. This study will serve to identify information gaps, and to provide evidence based recommendations to improve mortality statistics in Palestine.

METHODOLOGY

Study design

This retrospective death registry-based study examined a stratified random sample of DNFs of patients who died in hospitals in the West Bank and Gaza Strip and were reported in 2012. There are 81 hospitals in total in Palestine, with 51 in the West Bank (7 of them are in Jerusalem) and 30 in the Gaza Strip, serving around 4.78 million Palestinian (2.88 million in the West Bank and 1.9 million in Gaza Strip). Around 33% of hospitals in Palestine are governmental. The Ministry of Health accounts for 44% of the bed capacity in the West Bank, and 69% of the bed capacity in the Gaza Strip. We randomly selected 600 deceased from the Cause of Death Registry: 400 from the West Bank and 200 from the Gaza Strip. The sample size was estimated using the SampleXS program (2) and based on the following assumptions: the number of registered deaths in 2012 was 12,000; and the completeness of death certificates was 60% based

on the recent PCBS study of the completeness of the Cause of Death Registry (3); the margin of error 5%; and the design effect 1.5.

About 70% of the selected deaths in the West Bank occurred in government hospitals, while in the Gaza Strip all deaths were in government hospitals. We took a random sample of hospital deaths from the following four age groups: less than 1 year, 1-17 years, 18-64 years and 65+ years from Cause of Death Registry at PHIC. The study was carried out in 2014 in close collaboration with the Ministry of Health (MoH) in the West Bank and Gaza Strip. The response rate was 100%; all selected hospitals in the West Bank and Gaza Strip agreed to participate in the study.

Material

We used several sources of data to assess the accuracy and completeness of the mortality data (Figure 2). For each death in the sample we collected the following data:

(i) The Death Notification Form (DNF)

- Scanned image of the DNF
- Formal assessment of the DNF, such as legibility of the handwriting and if the administrative data had been filled out correctly with the certifier's signature.
- The causes of death in free text and administrative information were extracted from the DNF.

(ii) Data from PHIC:

- Date on which the death was registered.

- International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD 10) codes for underlying causes of death as recorded by the PHIC in the Death Registry.

(iii) Medical Extraction Forms (EDC):

Medical data on the death extracted from the deceased patient's hospital records. A team of specially trained physicians from each hospital in the study extracted data from the hospital record on the train of events leading to the death. The extraction file contained data on the reason for hospitalization, previous medical history and the course of events during hospitalization.

(iv) Hospital case summaries:

- Scanned image of the original hospital case summary.
- Case summary based on hospital records, completed by specially trained physicians in each hospital.

Figures 3 and 4 summarize data availability and attrition in the West Bank and Gaza Strip. In the West Bank, hospital records could be found for 371 cases of the 400 randomly selected deaths. In the Gaza Strip, we found hospital records for 199 of the 200 randomly selected deaths. We included all deaths reported in 2012 irrespective of date of death. The PHIC Registry gave an ICD code for the underlying cause of death in 358 of the 371 cases in the West Bank sample for which hospital records were available. In the Gaza Strip, 189 of the 199 cases in the sample had an ICD code for the underlying cause of death and 70 cases also included multiple causes of death registered by the PHIC. No multiple causes were present in the West Bank sample. Scanned images of the DNFs were available for 320 deaths in the West Bank

sample. EDC data on file were available for 365 West Bank deaths and for all Gaza Strip deaths. Table 1 summarizes the characteristics of the data.

Figure 2: Study Design Flow Chart

Table 1. Data characteristics of the records available for the study

Characteristic	N (%)	
	West Bank	**Gaza**
Sample size	**371**	**199**
- Government hospitals	265 (71%)	199(100%)
- Hospitals using electronic patient records	85 (23%)	0 (0%)
- Age		
< 1 year	75 (20%)	54 (27%)
1-17 years	45 (12%)	19 (10%)
18-64 years	149 (40%)	72 (36%)
65+	102 (27%)	54 (27%)
- Gender		
Male	200 (54%)	99 (50%)
Female	171 (46%)	100 (50%)

Figure 3: Data availability and attrition in West Bank

Figure 4: Data availability and attrition in Gaza Strip

STATISTICAL ANALYSIS

Completeness and timeliness

We did all analysis by region, due to the legislative and physical division of the occupied Palestinian territory, in terms of the separation of the Gaza Strip from the West Bank; resulting in two separate entities with different health systems and registries. We examined the proportion of complete DNFs in terms of personal data, immediate and underlying causes of death, and the attending physician and notifying person. We also examined the timeliness of registration, measured from the time of death to the time the deceased was registered in the Cause of Death Registry.

Underlying cause of death based on the patient's hospital file (EDC) and on the DNF

Using the EDC, we determined a "true" underlying cause of death based on the amended EDC dataset (please see below section, “Selecting the underlying cause of death”) and coded it using ICD-10. All coding, both of the original DNFs and of the EDC data, was done using Iris coding software (4). The version used (V4) corresponds to the 2013 edition of ICD-10. Iris fetches an ICD code for each diagnostic expression from a dictionary. Next, Iris prepares an input string of ICD codes for the ACME module (developed by the US National Center for Health Statistics). Finally, ACME selects an underlying cause according to the ICD rules and guidelines. After review and coding, we compared the three underlying causes – the one coded by PHIC; the one stated on the DNF form, but coded and classified by Iris; and the one according to the EDC and coded by researcher using IRIS– pair by pair:

- *Underlying cause on the DNF, coded by Iris vs Underlying cause in the PHIC Registry*

This comparison measures PHIC compliance with international rules for selection of the underlying cause of death.

- *Underlying cause on the DNF, coded by Iris vs Underlying cause according to the EDC*

This measures the accuracy of the original DNF.

- *Underlying cause according to the EDC vs Underlying cause in the PHIC Registry*

This comparison measures the overall accuracy of the underlying cause registered in the PHIC Registry.

We compared the underlying causes at two levels of detail: ICD detailed level (henceforth referred to as ICD 4-character level) and ICD block level. The principal use of measurements at detailed level data is to estimate the precision of the coding and classification, while measurements at ICD block level provide a broad assessment of the general trustworthiness for public health purposes.

For cross tabulation, we examined the accuracy of mortality statistics by gender and age group of the deceased patient, hospital affiliation (governmental and non-governmental), type of hospital records and underlying cause of death using chi square test. Tests of significance were two-sided with p- value ≤ 0.05. We divided the underlying causes according to the EDC into eight main diagnostic groups based on the frequency of the most common underlying causes according to the EDCs.

- Neoplasms (ICD-10 Chapter II)
- Metabolic diseases (ICD-10 Chapter IV)
- Cardiac diseases (ICD-10 I00-I51)
- Cerebrovascular diseases (ICD-10 I60-I69)

- Perinatal conditions (ICD-10 Chapter XVI)
- Congenital anomalies (ICD-10 Chapter XVII)
- External causes (ICD-10 Chapter XX)
- Other causes of death (ICD-10 chapters and codes not included elsewhere). ("Other" includes infectious diseases, diseases of the blood, neurological diseases, diseases affecting vision and hearing, respiratory diseases, gastrointestinal diseases, skin diseases, musculoskeletal diseases, urogenital diseases and symptoms with no clear connection to a single underlying disease.)

Selecting the underlying cause of death

The classification expert (LAJ) checked the diagnoses registered in the hospital patients' files against the DNFs and corrected the registered text if it differed from the DNF. Similarly, LAJ checked the causes of death recorded in the EDC dataset against the case summaries and against the scanned images of the hospital summaries. If an EDC entry was inconsistent with the case summary or the hospital summary, the text in the EDC dataset was corrected according to the summaries. The underlying causes of death derived from this amended EDC dataset were considered as the true underlying causes of death. .

Also, for each DNF, the following markers were used to record non-compliance with WHO instructions for completion of the death certificate: abbreviations used in the medical section, illegible writing in the medical section, sequencing errors in Part 1, symptomatic or secondary condition reported as underlying cause of death, cause of death insufficiently specified, and other reporting errors (such as reporting several causes on the lowest completed

line in Part 1, placing the underlying cause in Part 2 of the certificate, or using the wrong field on the DNF to report causes of death).

Patient and Public Involvement

Not applicable for this study as we used mortality statistics and data was de-identified prior to data analysis.

RESULTS

Completeness of DNF forms and other characteristics

Table 2 gives a summary of other characteristics of DNFs unrelated to the accuracy of the underlying cause of death. Although most deaths had been reported within a year (median of 89 days; range excluding extreme values 0-365 days), occasionally deaths from the 1980s and 1990s were only reported to the authorities in 2012. A fairly high number of DNFs (23% West Bank, 26% Gaza) reported an underlying cause of death that presumably developed as a complication of some other condition, yet the underlying condition is not recorded. For example, in several DNFs, kidney failure was reported as the underlying cause of death, but there was no mention of the reason why kidney failure developed, such as diabetes, glomerulonephritis or urinary obstruction. Administrative data (Part 1 of the DNF) were complete in only 7% and 2% of DNFs from hospitals with electronic patient records and paper patient records respectively. Other problems with electronic records included several deaths that had been registered twice or more, and some of the deceased recorded in the hospital files were still alive.

Table 2. Completeness and other DNF characteristics

	West Bank (N=320)		Gaza Strip (N=199)	
	cases	%	cases	%
No medical data on DNF	23	7.2	1	0.5

- electronic records (71)	15	21.1	-	-
- paper records (249)	8	3.2	1	0.5
Administrative data complete (Part I)	9	2.8	108	54.3
- electronic records (71)	5	7.0	-	-
- paper records (249)	4	1.6	-	-
Abbreviations used	132	41.2	76	38.2
Illegible writing	96	30.0	11	5.5
Sequence errors	64	20.0	38	19.1
Non-informative UC	74	23.1	51	25.6
UC lacking specificity	36	11.2	12	6.0
Irrelevant information	105	32.8	22	11.0
Other certification errors	92	28.8	40	20.1
Competing causes of death	36	11.2	23	11.6
Certifier's signature complete	159	59.7	50	25.1
Registry delay (days; median)	89		57	

Agreement between the three underlying causes (Table 3)

DNF- EDC agreement

Agreement on the underlying cause was highest between the underlying cause derived from the original DNF and the medical extraction forms (DNF-EDC). At the most detailed level, the agreement was 56% for the West Bank sample and 52% for the Gaza Strip sample. The difference between the two sample groups was not statistically significant (p-value: 0.37).

DNF-PHIC agreement

DNF-PHIC agreement between the DNF coded in line with international standards and the underlying cause as coded by the PHIC was 23% in the West Bank sample at ICD 4-character level and 39% in the Gaza sample. The difference between the samples was statistically significant (p-value: 0.0001). At ICD block level, the West Bank sample had a DNF-PHIC agreement of 46% and the Gaza sample 53%, a difference that was not statistically significant (p-value:0.12). .

EDC-PHIC agreement

Agreement was lowest between the EDC and the underlying cause recorded in the PHIC Registry. At ICD 4-character level, EDC-PHIC agreement was 19% in the West Bank sample and 31% in the Gaza sample.

Table 3. Agreement on underlying cause by ICD level: point estimates, and for ICD-4 character level and ICD block level and 95% confidence intervals

		ICD 4-character		ICD block level	
		PE (%)	95% CI	PE (%)	95% CI
DNF-EDC	West Bank	56	51-62	68	63-73
	Gaza Strip	52	44-59	62	55-68
DNF-PHIC	West Bank	23	19-28	46	41-51
	Gaza Strip	39	31-46	53	46-60
EDC-PHIC	West Bank	19	15-23	44	38-49
	Gaza Strip	31	24-38	44	37-51

Agreement by sex, age, type of hospital and type of deceased hospital records

In both samples, there was no noticeable relationship between the sex and age of the decedent and the accuracy of the cause of death There were no statistically significant differences between males and females in underlying cause agreement (21% vs.16%, respectively, p-value: 0.16). Also, there were few significant differences between the four age groups, either at ICD 4-character or ICD block level. At a detailed level there was, significantly lower agreement for the youngest age groups in the two comparisons involving the codes in the PHIC Registry (DNF-PHIC and EDC-PHIC) (7% and 8%) compared to 27% and 24% at age group 18-64years (both p-value: <0.0001). At block level, only the DNF-PHIC comparison showed a significant difference between children <1 year and older people >64 (43% vs 60% respectively, p-value: 0.0002), probably reflecting difficulties in the coding of congenital anomalies and perinatal causes.

Also, the West Bank sample showed only minor differences between governmental and non-governmental hospitals, and no differences in accuracy between DNFs from hospitals using an electronic patient record system or those using paper-based documentation. The only statistically significant difference between governmental and non-governmental hospitals was in DNF-EDC agreement at ICD 4-character level (53% and 64% respectively, p-value: 0.013), but this difference was not significant at ICD block level (data not shown). With the exception of EDC-PHIC comparison at a detailed level (11% using electronic records and 21% using paper records, p-value: 0.0003), the use of an electronic patient record system or traditional paper medical records by the hospital had no impact on agreement. The EDC-PHIC difference disappeared at block level.

Agreement by diagnostic group (Table 4)

DNF- EDC agreement

There were considerable differences in agreement between diagnostic groups. At both the ICD 4-character level and block level, the accuracy of the DNFs (DNF-EDC agreement) was best for external causes and neoplasms. The weakest agreement was for metabolic conditions. The differences between the West Bank and Gaza samples were not statistically significant (56%, 52% respectively, p-value:0.37).

DNF- PHIC agreement

At both ICD detailed level and ICD block level, agreement in the West Bank sample between the original DNFs and PHIC Registry data was best for cardiac diseases, but still weak. It was generally better in the Gaza sample, with the highest value for neoplasms (39% , 64% respectively, p-value: <0.0001) and cerebrovascular conditions (16% , 63% respectively, p-value: <0.0001), both significantly better than in the West Bank sample. The lowest DNF-PHIC agreement in the West Bank sample was for anomalies and external causes significantly lower than in the Gaza sample (3%, 29% respectively, p-value:<0.0001). In the Gaza sample, the lowest DNF-PHIC agreement was for perinatal causes.

EDC-PHIC agreement

The accuracy of the underlying causes in the PHIC register, measured by EDC-PHIC agreement, was highest, but still moderate, for neoplasms and cardiac conditions in the West Bank sample. In the Gaza sample the highest agreement was for neoplasms (62%, significantly higher than in

the West Bank sample 36%, p-value :< 0.0001). The lowest agreement was for metabolic conditions in the West Bank sample and for perinatal causes in the Gaza sample.

Table 4. Agreement in underlying cause by diagnostic group: ICD 4-character level, point estimates and 95% confidence intervals

	DNF-EDC				DNF-PHIC			
	West Bank		Gaza Strip		West Bank		Gaza Strip	
	PE (%)	95% CI	PE (%)	95% CI	PE (%)	95% CI	PE (%)	95% CI
Neoplasm	76	65-87	84	73-95	39	26-51	64	49-79
Metabolic	32	12-52	29	8-50	16	1-31	19	1-37
Cardiac	49	38-61	37	13-61	42	30-54	29	5-54
Cerebro-vascular	53	38-68	45	21-69	16	5-27	63	39-87
Perinatal	46	27-66	45	23-68	11	0-23	5	0-15
Anomaly	51	34-69	47	21-74	3	0-9	29	5-54
External	78	60-96	-		9	0-22	-	
Other	56	43-69	45	32-59	16	6-26	36	22-50

	EDC-PHIC			
	West Bank		Gaza Strip	
	PE (%)	95% CI	PE (%)	95% CI
Neoplasm	36	26-50	62	47-77
Metabolic	4	0-11	10	0-23
Cardiac	36	25-48	24	1-46
Cerebrovascular	7	0-15	26	5-48
Perinatal	11	0-21	5	0-15
Anomaly	3	0-9	24	1-46
External	9	0-22	-	
Other	10	2-18	31	18-45

DISCUSSION

This is the first baseline study of accuracy of mortality statistics in Palestine. Based on the study findings, the accuracy is low, mostly due to inaccuracies in the PHIC coding of the underlying cause of death. But deficient certification of causes of death also contributed significantly to the inaccuracy. Both the West Bank and Gaza, 30%-40% of DNFs had sequence-related errors combined with incorrect order or entries and competing causes of death. From an international perspective, this is a high figure (5). Agreement between the underlying cause according to the EDC and the underlying cause of death actually registered by the PHIC was low: 19% (West

Bank) and 31% (Gaza) at the most detailed level and 44% (both samples) even at ICD block level. In some comparisons the Gaza strip performed better than the West Bank, in others it was the other way round. Further research is needed to understand the differences in accuracy between Gaza Strip and West Bank

Based on study findings, the accuracy of mortality statistics is lower than in other published similar studies. For example, a validation of mortality statistics in Cape Town showed accuracy of 55% at WHO tabulation list 1 level (103 groups, roughly corresponding to block level) (6). Other studies vary widely in their assessment of the accuracy of mortality statistics, (7) but since they generally refer to mortality statistics in high-resource countries and to specific causes of death, the results are difficult to compare to the present study.

As in several other studies on the quality of mortality statistics, accuracy varies between diagnostic groups (7-9). Accuracy in the causes of death in the PHIC register, indicated by EDC-PHIC agreement, was best, but still moderate, at 69% for cerebrovascular diseases in the West Bank sample. This is surprising since most studies show the highest agreement for neoplasms, which is also the case in the Gaza sample (83%) (9). In the West Bank sample, neoplasms show 54% of agreement, which is low in comparison with other studies. Both samples show low agreement for metabolic conditions, which is consistent with findings from other studies.

Although the present study adds to the accuracy of mortality statistics registries' literature, it has several limitations. First, the study was limited to hospital deaths. A future study is needed to examine the accuracy of DNFs for deaths outside the hospital context. Second, pairwise percent agreement used in data analysis does not take into account that agreement may occur by chance. Therefore, the percent agreement may be over reported. Third,

we do not know the number of deaths that are not reported to the authorities at all. The presence of deaths that occurred several decades ago in the sample indicates that some deaths are unreported. It is important to estimate the number of non-reported deaths and determine if the causes of death in such cases differ significantly from those actually reported to the authorities.

CONCLUSIONS

Based on the study, the PHIC Cause of Death Registry gives a poor picture of the causes of death in both the West Bank and Gaza Strip. This hampers the assessment of important aspects of public health, such as the burden of disease, maternal mortality and the prevalence of congenital anomalies. The conclusion is that PHIC mortality statistics are not comparable to statistics from other countries that adhere to the ICD instructions more closely, and also that the medical precision of the statistics (measured by EDC-PHIC agreement) would be improved if the international coding and classification rules were adhered to. To address these points, the following measures are proposed:

1) Use of international coding software for coding and classification (10).

The use of internationally recognized coding software would bring several benefits. The software automatically applies ICD instructions for the selection and classification of causes of death. It also covers less common cases that coders may be unfamiliar with and brings coding and classification in line with the ICD instructions. A dictionary of medical terms is included in the software, contributing further to consistent coding. Further, coding software speeds up the coding process since it reduces the burden on the limited staff at PHIC who are currently responsible for all the coding. At present there can be a delay of several months from the time the DNF is received by the PHIC until it is coded and entered in the CoDR. A further advantage of automated coding is that with automated coding, the PHIC Registry would contain multiple

causes of death for all deaths, and the potential for detailed analysis and monitoring would be considerably enhanced.

2) Training and supervision of cause of death certification at hospitals

Studies of the quality of mortality statistics often suggest that training for certifiers on how to complete the death certificate would resolve issues related to quality. An automated coding system would facilitate the training of new coders. Another proposed approach is that each death certificate should be written and signed by two physicians (11, 12), or that a senior physician should review all death certificates issued at the hospital. In some locations, a specialist committee has been set up to review all hospitalizations that end with the death of the patient, and also to verify the accuracy of the death certificate (12, 13).

3) Feedback from PHIC to the individual hospital on problems listed in Table 2 (14,15). It would be even more efficient if the hospital or certifier could be contacted in such cases and asked to amend the DNF.

In addition, steps should also be taken to improve the overall completeness of the Cause of Death Registry. First, review and improve data handling at hospitals with electronic patient records. Electronic patient record systems are supposed to improve data availability and facilitate the administration of hospital care, but appeared to have the opposite result in our study. Second, develop a digital DNF in the electronic system. Second, enforce completion of the death report in the electronic system by adding specific restrictions in the electronic system.

According to a study conducted by the Palestinian Central Bureau of Statistics, it now includes only about 60% of all deaths (3). Consequently, the statistics would still provide a poor picture of the causes of death even if the accuracy of reported cases was considerably improved.

Acknowledgements The authors thank M Abu Ghali, J Bitar, U Balawi, S Khammash, D Ghoushe, M Ali and E Aly.

Contributors LA conceptualised and designed the study. SM and IR were responsible for data collection. LA, SM and HD were responsible for analysis of the data. SM, LA, HD, IR, RS, AR contributed to the preparation of the manuscript, revised, and approved the final draft.

Competing interest: None declared

Ethical approval The Helsinki Committee for Ethical Approval- Palestinian Health Research Council, Gaza Strip/Palestine approved the study protocol in January 2014 (Ethical approval number: PHRC/HC/67/14).

Data sharing statement No additional data available.

Funding This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

1. World Health Organization. Strengthening civil registration and vital statistics for births, deaths and causes of death: resource kit. 2013.
2. Sample XS. http://external.informer.com/brixtonhealth.com/samplexs.html
3. Palestinian Central Bureau of Statistics. Assessment of birth and mortality data in the West Bank 2011-2012. Ramallah: Palestinian Central Bureau of Statistics, 2013.
4. German Institute of Medical Documentation and Information. Iris Institute. Germany 2015 [cited 2016 October 18, 2016]; Available from: http://www.dimdi.de/static/en/klassi/irisinstitute/index.htm.
5. Jougla E, Rossolin F, Niyonsenga A, Chappert J, Johansson L, Pavillon G. Comparability and quality improvement in European causes of death statistics. Eurostat project. 2001;96:S99-5761.
6. Burger EH, Groenewald P, Bradshaw D, Ward AM, Yudkin PL, Volmink J. Validation study of cause of death statistics in Cape Town, South Africa, found poor agreement. Journal of clinical epidemiology. 2012;65(3):309-16.
7. Johansson LA, Westerling R, Rosenberg HM. Methodology of studies evaluating death certificate accuracy were flawed. Journal of clinical epidemiology. 2006;59(2):125-31.
8. Johansson LA, Björkenstam C, Westerling R. Unexplained differences between hospital and mortality data indicated mistakes in death certification: an investigation of 1,094 deaths in Sweden during 1995. Journal of clinical epidemiology. 2009;62(11):1202-9.
9. Johansson LA. Targeting non-obvious errors in death certificates. Doctoral dissertation, Acta Universitatis Upsaliensis. 2008.

10. German Institute of Medical Documentation and Information. Iris Institute. Secondary Iris Institute 2018. https://www.dimdi.de/dynamic/en/classifications/iris-institute/.

11. Lahti RA, Penttilä A. Cause-of-death query in validation of death certification by expert panel; effects on mortality statistics in Finland, 1995. Forensic science international 2003;131(2-3):113-24.

12. Hanzlick R. Quality assurance review of death certificates: a pilot study. The American journal of forensic medicine and pathology 2005;26(1):63-65.

13. Curb JD, Davis BR, Tung B. Standardized physician preparation of death certificates. Controlled Clinical Trials 1985;6(3):237.

14. Jougla E. Comparability and Quality Improvement of European Causes of Death Statistics. European Commission DG Sanco Agreement. Final report July 2001. EDC DGV/F3 SOC 98 20108.

15. Hoyert DL, Lima AR. Querying of death certificates in the United States. Public Health Rep. 2005;120(3):288-93.

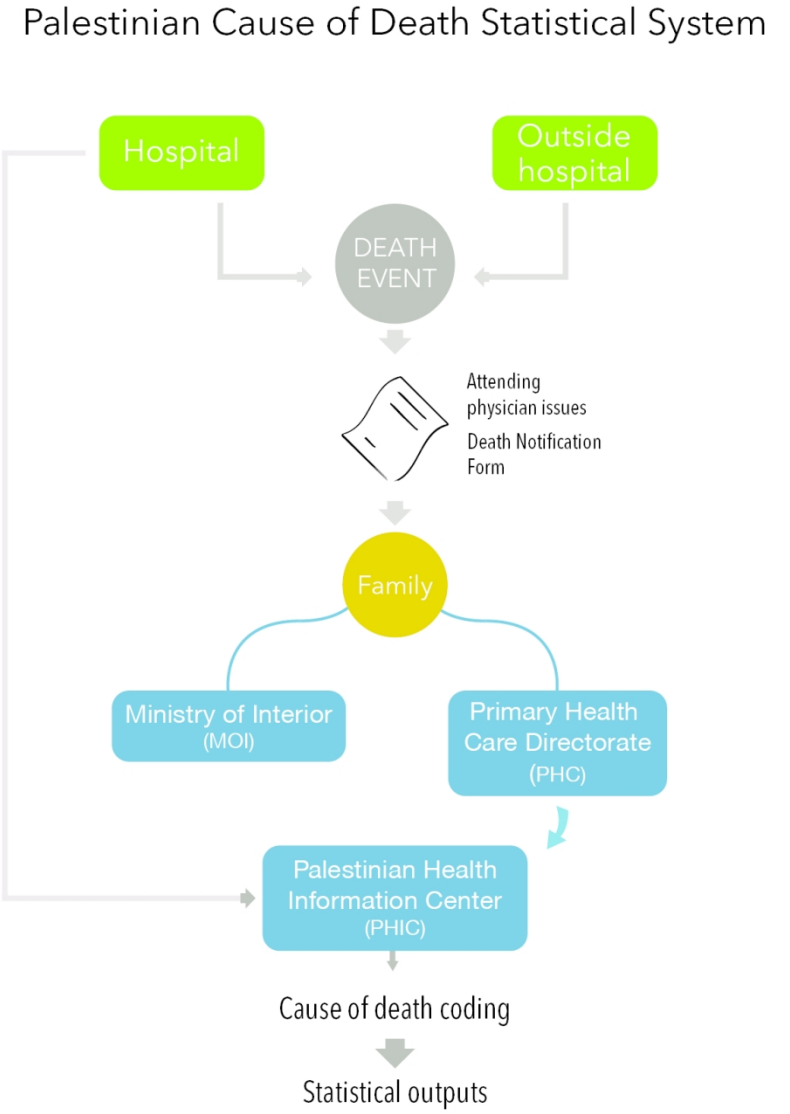


Figure 1: Palestinian Cause of Death Statistical System

209x297mm (150 x 150 DPI)

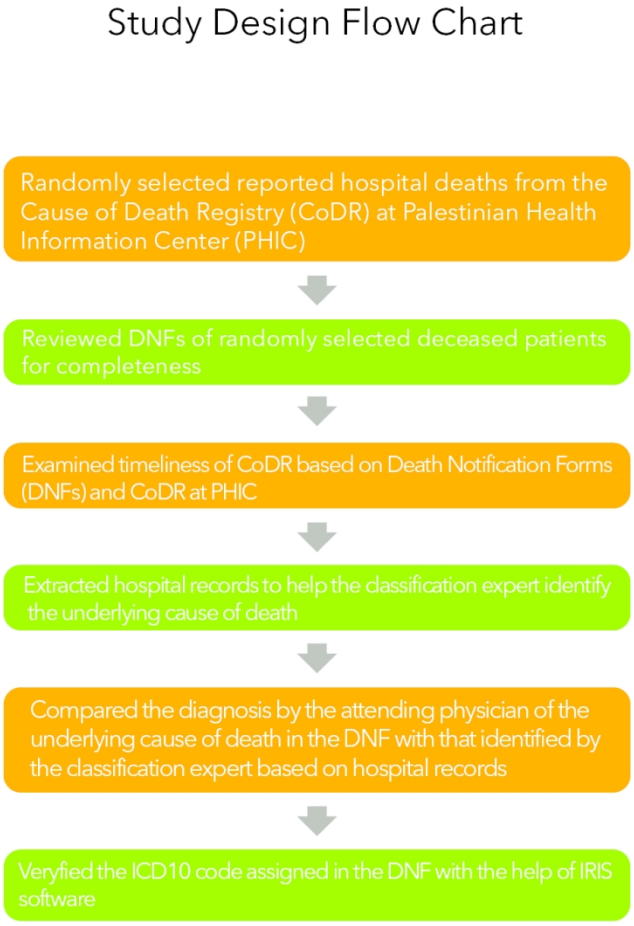


Figure 2: Study Design Flow Chart

209x297mm (150 x 150 DPI)

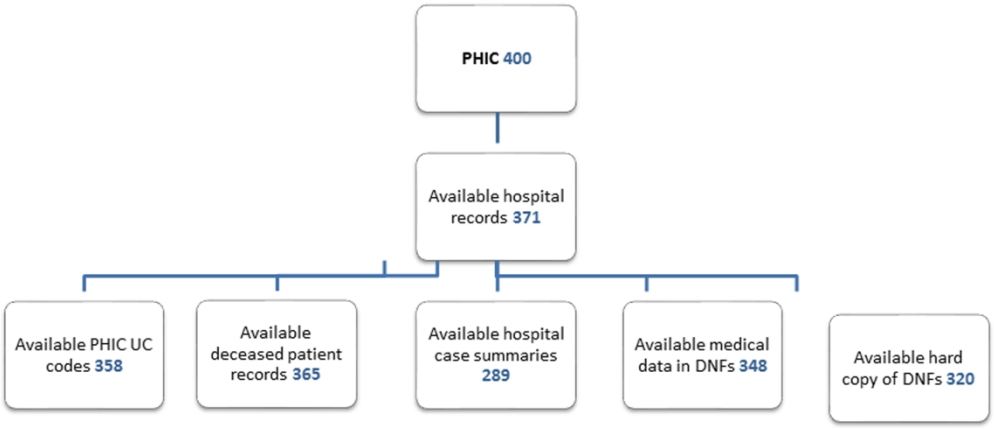


Figure 3: Data availability and attrition in West Bank

297x155mm (300 x 300 DPI)

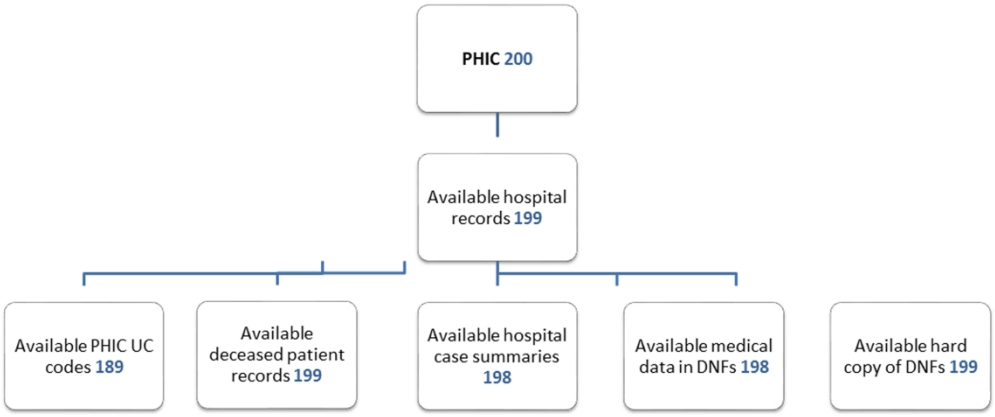


Figure 4: Data availability and attrition in Gaza Strip

297x155mm (300 x 300 DPI)

STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of *cross-sectional studies*

Section/Topic	Item #	Recommendation	Reported on page #
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	3
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	5-6
Objectives	3	State specific objectives, including any prespecified hypotheses	5
Methods			
Study design	4	Present key elements of study design early in the paper	6-8
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6-10
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants	6-10
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	7-8, 11-13
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	7-8, 11-13
Bias	9	Describe any efforts to address potential sources of bias	*13*
Study size	10	Explain how the study size was arrived at	6
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	11-13
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	11-13
		(*b*) Describe any methods used to examine subgroups and interactions	12-13
		(*c*) Explain how missing data were addressed	7-10
		(*d*) If applicable, describe analytical methods taking account of sampling strategy	11-13
		(*e*) Describe any sensitivity analyses	
Results			

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	7-8, 10
		(b) Give reasons for non-participation at each stage	NA
		(c) Consider use of a flow diagram	10
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	9
		(b) Indicate number of participants with missing data for each variable of interest	9-10
Outcome data	15*	Report numbers of outcome events or summary measures	14-19
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	14,16.19
		(*b*) Report category boundaries when continuous variables were categorized	16-19
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	17
Discussion			
Key results	18	Summarise key results with reference to study objectives	19-20
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	20
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	21-23
Generalisability	21	Discuss the generalisability (external validity) of the study results	20
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	23

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.